University of Alberta

MEASUREMENT OF WALKING CAPACITY IN LUMBAR SPINAL STENOSIS

by

Christy Tomkins

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

REHABILITATION SCIENCE

FACULTY OF REHABILITATION MEDICINE

Edmonton, Alberta
FALL 2008

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-46438-0
Our file Notre référence
ISBN: 978-0-494-46438-0

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

For Julie

She was an inspiration to us all.

ABSTRACT

Background

Measurement of walking capacity in lumbar spinal stenosis (LSS) is important for both research and clinical practice. Measures that have been used for this purpose include the Physical Function Scale, the Oswestry Disability Index (ODI) and treadmill protocols. However, it is not known whether the construct being measured using these tests is in fact walking capacity. To date, no criterion measure of walking capacity has been established for this population. For the purpose of this research, the Self Paced Walking Test (SPWT) was developed for use as a criterion measure, based on an operational definition of the construct of walking capacity.

Purpose

The aim of this research was to provide validity evidence regarding the use of a treadmill protocol and various self-report measures of walking capacity in LSS.

Methods

In a preliminary study, data from a prospective study of LSS was used to investigate validity of the Physical Function Scale for the measurement of walking capacity (n=72). A second sample of LSS patients was then recruited to examine both construct and criterion validity evidence regarding the use of a treadmill protocol and various self-report measures of walking capacity, using the SPWT as the criterion measure (n=41). Reproducibility of the SPWT and various self-report instruments was also examined using a subgroup of this second sample (n=28). All subjects included in this research had LSS confirmed on imaging and

by a spine specialist surgeon. Correlational analyses were employed for all studies.

Results and Conclusions

Validity evidence was provided supporting use of the treadmill test, yet it was concluded that this test significantly underestimates patients' walking capacities. Evidence of validity and reproducibility was provided supporting the use of the condition specific Physical Function Scale. While the ODI as a whole was not found to be an appropriate measure of walking capacity, its walking distance specific item was found to be both reproducible and highly correlated with the criterion. In addition, validity and reproducibility evidence confirmed the utility of the SPWT as the criterion measure. Results of this research provide valuable information regarding the selection and interpretation of measures of walking capacity for use with LSS patients.

ACKNOWLEDGEMENTS

I wish to thank my doctoral supervisor, Dr. Michele Crites Battié for her continual support, patience and guidance throughout my program. Her knowledge and expertise were invaluable. I could not have asked for a better mentor. I am going to miss having the luxury of dropping in to talk to Michele, and hope we can maintain a relationship in years to come.

I would also like to express gratitude to my supervisory committee members, Dr. Todd Rogers and Dr. Stuart Petersen for their consistent feedback and participation throughout my program. Dr. Rogers was an unbelievable asset for advice regarding methodological issues, data collection and statistical analysis. I enjoyed our many discussions and debates. Dr. Petersen was instrumental helping to organize and facilitate my research plans.

I would like to acknowledge Angela Libutti for her contributions to keeping my program organized. I wish also to thank the administrative staff and spine surgeons involved with recruiting subjects for my research. Maureen Tooke, the assistant to Dr. Harry Jiang was instrumental in providing me with research subjects. I am grateful for the financial support for both my doctoral studies and research, provided by the Alberta Canadian Institutes for Health Research Training Program in Bone and Joint Health, the Alberta Heritage Foundation for Medical Research, the University of Alberta Faculty of Graduate Studies and Small Faculties Granting Program.

Lastly, and most importantly, I would like to thank my family and friends for their unwavering support, love and patience throughout my studies. Haydn has been a great source of motivation and a wonderful listener. Amber has been my greatest cheerleader. My Aunt Virginia and Grandmother Barbara have been a source of love and support from the beginning, while my cousin Lee has been an invaluable academic resource. My greatest source of support has always been my parents, Barb and Jim. They have been there for me in every way and have supported all of my decisions with enthusiasm. They have never doubted me for a second, which is the greatest gift anyone can ask for. I could not have done this without them.

TABLE OF CONTENTS

Page

Chapter 1: Introduction..... 1
1.1 Overview..... 1
1.2 Purpose..... 4
1.3 References..... 5

Chapter 2: Background.....8
2.1 Lumbar Spinal Stenosis Definition and Clinical Presentation.....8
2.1.1 Neurogenic Claudication.....9
2.2 Diagnosis of LSS.....9
2.2.1 Historical and Physical Findings in LSS Patients.....9
2.2.2 Imaging..... 10
2.3 Patho-Anatomy of LSS..... 11
2.4 Pathophysiology of Neurogenic Claudication..... 12
2.4.1 The Ischemic Theory of Neurogenic Claudication.....12
2.4.2 The Venous Stasis Theory of Neurogenic Claudication.....13
2.4.3 Posture and Neurogenic Claudication.....13
2.5 Natural History of LSS.....14
2.6 Walking Capacity in LSS.....14
2.6.1 Definition of the Construct of Walking Capacity in LSS..... 15
2.6.2 Measures of Walking Capacity for Outcomes of Treatment and Diagnosis..... 15
2.7 Treadmill Measurement of Walking Capacity in LSS..... 16
2.7.1. Treadmill Protocols for which there Exists Psychometric Evidence.....17
2.7.2 Examining the Diagnostic Value of Treadmill Testing in LSS 20
2.7.2 Other Observational Measures of Walking Capacity in LSS... 20
2.8 Self-report Measures of Walking Capacity in LSS.....21
2.8.1 Physical Function Scale of the Swiss Spinal Stenosis Questionnaire.....22
2.8.2 Oswestry Disability Index.....24
2.8.3 Health Utilities Index Mark 3 Ambulation Score.....24
2.8.4 Visual Analog Scale for Walking Difficulty..... 25
2.9 Summary..... 25
2.10 References..... 28

Chapter 3: Validity..... 35
3.1 Overview.....35
3.2 Building a Construct Validity Argument.....35
3.3 Types of Validity Evidence.....36
3.3.1 Content Validity Evidence.....36

3.3.2 Criterion Validity Evidence.....37
3.4 Validity as it Applies to the Current Thesis Research.....38
3.5 References.....40

Chapter 4: Study 1 - Construct Validity of the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire for the Measurement of Walking Capacity..... 41
4.1 Introduction..... 41
4.2 Materials and Methods.....44
4.2.1 Subjects.....44
4.2.2 Data Collection.....44
4.2.3 Hypotheses.....45
4.2.4 Data Analysis.....46
4.3 Results..... 46
4.4 Discussion.....47
4.5 Conclusions.....50
4.6 References.....54

Chapter 5: Study 2 - Reproducibility of Measures of Walking Capacity in Lumbar Spinal Stenosis..... 57
5.1 Introduction..... 57
5.2 Materials and Methods..... 59
5.2.1 Subjects..... 59
5.2.2 Data Collection.....59
5.2.3 Self Paced Walking Test Protocol..... 60
5.2.4 Data Analysis.....61
5.3 Results.....62
5.3.1 SPWT Total Walking Distance.....63
5.3.2 SPWT Distance to First Symptoms and Walking Speed..... 64
5.3.3 Self-Report Measures.....64
5.3.4 Change Score..... 64
5.3.5 Post-hoc Analyses.....65
5.4 Discussion.....66
5.4.1 Self-Paced Walking Test as a Criterion Measure of Walking Capacity in LSS..... 66
5.4.2 Total Walking Distance in the SPWT..... 67
5.4.3 Distance to First Symptoms.....71
5.4.4 Walking Speed.....71
5.4.5 Physical Function Scale..... 71
5.4.6 Oswestry Disability Index..... 72
5.4.7 Construct Specific Distance Question and Walking Distance Specific Items..... 73
5.4.8 Can we Generalize Results of the SPWT Reproducibility to all LSS Patients?.....73
5.4.9 Change Scores.....75
5.4.10 Limitations.....76

5.5 Conclusions..... 76
5.6 References..... 81

Chapter 6: Study 3 - Validity of Measures of Walking Capacity in Lumbar Spinal Stenosis..... 84
6.1 Introduction..... 84
6.1.1 Treadmill Testing of Walking Capacity in LSS..... 84
6.1.2 Self-report Measures of Walking Capacity in LSS..... 85
6.2 Materials and Methods..... 88
6.2.1 Subjects..... 88
6.2.2 Data Collection..... 88
6.2.3 Criterion Measure: Self-Paced Walking Test..... 89
6.2.4 Treadmill Protocol..... 89
6.2.5 Questionnaire..... 90
6.2.6 Hypotheses for Convergent and Divergent Construct Validity..... 91
6.2.7 Hypotheses for Criterion Validity..... 92
6.2.8 Data Analysis..... 92
6.3 Results..... 93
6.3.1 Subjects..... 93
6.3.2 Walking Tests..... 93
6.3.3 Convergent and Divergent Construct Validity Evidence..... 93
6.3.4 Criterion Validity Evidence: Treadmill..... 94
6.3.5 Criterion Validity Evidence: Self-Report Measures..... 95
6.3.6 Post-hoc Analyses..... 95
6.4 Discussion..... 96
6.4.1 The SPWT as the Criterion Measure of Walking Capacity in LSS..... 96
6.4.2 Validity of Treadmill Measurements..... 97
6.4.3 Validity of Self-Report Instruments and Items..... 99
6.4.4 Physical Function Scale..... 100
6.4.5 Oswestry Disability Index..... 101
6.4.6 Self-Reported Walking Distance..... 102
6.4.7 Can we Generalize Validity Results to all LSS Patients?..... 102
6.4.8 Limitations..... 103
6.5 Conclusions..... 104
6.6 References..... 107

Chapter 7: General Discussion and Conclusions..... 112
7.1 Overview..... 112
7.2 The Validity Argument..... 112
7.3 The SPWT as the Criterion Measure of Walking Capacity in LSS..... 113
7.3.1 Rationale for the Choice of Criterion Measure Used in the Present Research..... 114
7.3.2 Choice of Speed for the Criterion..... 114
7.3.3 Choice of Track vs. Treadmill as the Criterion..... 114

7.3.4 Choice of Grade for the Criterion..... 115
7.3.5 Test Duration of the Criterion..... 115
7.3.6 Reproducibility of the Criterion and Stability of Walking in LSS Patients.....116
7.3.7 Validity of the Criterion (SPWT)..... 118
7.4 Treadmill Testing of Walking Capacity in LSS..... 119
7.4.1 Rationale for Choice of Treadmill Protocol.....119
7.4.2 Validity of the Treadmill Test..... 119
7.5 Physical Function Scale of the Swiss Spinal Stenosis Questionnaire.....121
7.5.1 Reproducibility of the Physical Function Scale.....121
7.5.2 Validity of the Physical Function Scale.....122
7.6 Oswestry Disability Index.....123
7.6.1 Reproducibility of the Oswestry Disability Index..... 123
7.6.2 Validity of the Oswestry Disability Index..... 123
7.7 Self-Reported Walking Distance Reproducibility and Validity..... 124
7.8 Walking difficulty Scale Reproducibility and Validity.....125
7.9 Characteristics of the Study Samples.....125
7.10 Limitations.....127
7.11 Future Research..... 129
7.11.1 Reproducibility of Treadmill Testing in LSS Patients.....129
7.11.2 Validity and Reproducibility of Measures of Walking in LSS Subgroups..... 130
7.11.3 Predictors of Walking Capacity in LSS..... 130
7.11.4 Responsiveness of Measures of Walking in LSS.....131
7.11.5 Community-Based Measurement of Walking Capacity..... 131
7.12 Conclusions 132
7.12 References 135

Appendix A: Forms.....140

Appendix B: Detailed Methodology Regarding Data Collection For Chapters 5 and 6 (Studies 2 and 3)..... 162

LIST OF TABLES

Page

Table 2-1. Treadmill Protocols used to Measure Walking in LSS..... 27

Table 4-1. Psychometric Properties of the Physical Function (PF) Scale of the Swiss Spinal Stenosis Questionnaire..... 52

Table 4-2. Characteristics of the Study Group (n=72)..... 53

Table 4-3. Convergent and Divergent Correlation Coefficients for the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire in the Evaluation of Walking in Subjects with Lumbar Spinal Stenosis (n=72)..... 54

Table 5-1. Individual Differences for Repeat Administrations: SPWT Total Walking Distance..... 76

Table 5-2. The Difference between Test Re-test Administrations of Walking Capacity Measures in Patients with LSS (n=28)..... 77

Table 5-3. ICC and Confidence Intervals for Test Re-test Reproducibility of Measures of Walking in LSS (n=28)..... 78

Table 5-4. The Difference between Test Re-Test Administrations of Walking Capacity Measures in Patients with LSS who Reported no Change between Tests (n=15)..... 80

Table 6-1. Walking Capacity Measures (n=41)..... 103

Table 6-2. Convergent and Divergent Validity Coefficients among Measures of Walking Capacity in Lumbar Spinal Stenosis (n=41)..... 104

Table 6-3. Individual Differences in Distance between Self-Paced Walking Test and Treadmill Test..... 105

Table 6-4. Convergent and Divergent Validity Coefficients among Measures of Walking Capacity in Lumbar Spinal Stenosis Excluding Subjects who Walked a Full 30 Minutes in the SPWT (n=33)..... 106

Table 7-1. Characteristics of the Study Samples..... 133

CHAPTER 1
INTRODUCTION

1.1 Overview

Degenerative lumbar spinal stenosis (LSS) is a debilitating condition which typically affects adults in their sixth and seventh decades of life, and is often a cause of chronic low back and leg pain in older adults.[1] As the general population ages, LSS is being encountered more frequently.[2;3] Recent advances in imaging technology and improvement in diagnostic accuracy have also contributed to this marked increase in the diagnosis of LSS internationally.[3-5] LSS has become one of the conditions seen most frequently in orthopaedic and neurosurgery practice.[6] It has been reported that in the United States, LSS is now the most common diagnosis among patients over sixty undergoing lumbar spine surgery.[7] Deyo et al.[12] suggest that this trend toward increasing numbers of surgical procedures for LSS has continued to grow into the 21st century, and will no doubt continue to grow with the aging population. As such, LSS is, and will continue to be, associated with significant healthcare costs.[8-10] It has been estimated that the total inpatient expense of LSS exceeds 1 billion dollars annually in the US.[8; 9; 11] Unfortunately, similar data regarding the health care cost of LSS in Canada is not available. However, the trend toward increasing numbers of LSS diagnoses is internationally recognized.[3] Given the significant economic ramifications associated with treatment for this increasingly prevalent diagnosis, the identification of effective treatment options and psychometrically sound assessment methods for this population is a priority.[13]

Anatomically, LSS refers to a narrowing of the central spinal canal, lateral recesses and/or intervertebral foramen causing compression of the associated neurovascular structures. Degenerative lumbar stenosis results from changes of the spine that occur with aging, including facet joint hypertrophy, loss of intervertebral disc height, disc bulging, osteophyte formation and hypertrophy of the ligamentum flavum.[14] The most specific symptom of degenerative central LSS is neurogenic claudication, which causes pain, numbness, weakness and tingling

in the low back, buttocks and legs during walking or standing. Accordingly, due to pain and discomfort in the lower extremities, people with LSS often avoid walking and have reduced walking capacities.

Measurement of walking capacity is very important in people with LSS given the clinical presentation of the condition. Clinicians and researchers often measure walking when assessing treatment outcomes, physical function and the natural progression of the condition. Deen et al. suggest that there has been a recent trend toward emphasizing walking capacity as a key outcome indicator for patients with LSS symptoms,[15] while, similarly, Yamashita et al. found self-reported difficulty with walking to be the strongest and only independent predictor of patient satisfaction post-surgery for LSS.[16] Thus, with an aging population and increasingly prevalent diagnosis of LSS, there is demand for valid and reproducible tests of walking for use with LSS patients.[15;16]

For the purpose of this thesis research, the construct of *walking capacity* was defined as the distance a person with LSS is able to walk without support, on a level surface at a self-selected speed before being forced to stop due to symptoms of LSS. To date there is no 'gold standard' or criterion measure of walking capacity in this population. Instead, there are a number of self-report and observational measures used in research with LSS populations which address to a certain degree the construct of walking capacity as defined here. These measures include the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire, [17-24] the Oswestry Disability Index (ODI) [25-33] and various treadmill protocols.[32;34-41]

Treadmill testing of patients with LSS has been shown to be safe, inexpensive, quantifiable and easily administered.[15] However, limited validity evidence has been provided regarding the use of treadmill tests for evaluating walking in LSS. Of the self-report instruments used in LSS research, both the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire [17-24] and the low back specific ODI [25-33] have been used to assess walking capacity in people with LSS. The Physical Function Scale was designed specifically for use in the evaluation of walking in patients with LSS,[1] and has been used in a number of

studies for this purpose.[25-33] While there is some literature focusing on the psychometric properties of the Physical Function Scale, this literature is limited.[1;41;42] None of the existing studies were designed to specifically assess the validity of the Physical Function Scale for the measurement of walking capacity in persons with LSS. The ODI is a well-validated instrument for use in the evaluation of lumbar conditions,[1;39;41;43-47] yet it does not address the specific neuro-ischemic characteristics of LSS (lower limb numbness, weakness and tingling). No studies have examined the validity of the ODI for measurement of walking capacity in LSS populations.

Given that there is no gold standard measure available to tap the construct of walking capacity in LSS, an observational walking test was developed for use as a gold standard or criterion measure in this thesis research. This measure, the 'Self Paced Walking Test' (SPWT), is intended to be as relevant to and representative of the defined construct of walking capacity in LSS as possible. The SPWT requires patients to walk around a track without support and at their own pace until forced to stop due to symptoms of LSS (or a time limit of 30 minutes). The SPWT is meant to mimic authentic walking conditions using a standardized setting and protocol.

Unfortunately, the SPWT may be impractical in many research and clinical settings, given that not all facilities have access to a track. It would be possible to use a treadmill protocol as a surrogate of this criterion measure. However, not all clinicians and researchers have access to a treadmill either. Realistically, in most settings, the use of self-report measures of walking capacity may be most feasible. Yet, very little research has been reported examining the psychometric properties of any of the currently employed measures of walking capacity in LSS. It is not known how treadmill or self-report based measures of walking capacity are associated with a criterion measure of walking, or whether there is validity evidence regarding the use of these measures as surrogates of a criterion in the assessment of walking capacity in persons with LSS.

1.2 Purpose

Thus, the purpose of this thesis research was to examine validity evidence regarding the use of various measures of walking capacity in lumbar spinal stenosis. The specific objectives included:

1. Providing construct validity evidence regarding the use of the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire for the measurement of walking capacity in persons with lumbar spinal stenosis.
2. Examining the reproducibility of the criterion measure of walking capacity (the Self-Paced Walking Test), as well as various self-report measures of walking.
3. Examining the validity evidence regarding the use of a treadmill protocol and various self-report instruments as surrogate measures of walking capacity in LSS, using the Self-Paced Walking Test as the criterion measure.

1.3 References

1. Stucki G, Daltroy L, Liang MH et al. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine.* 1996;21:796-803.

2. Haig AJ, Tong HC, Yamakawa KS et al. Spinal stenosis, back pain, or no symptoms at all? A masked study comparing radiologic and electrodiagnostic diagnoses to the clinical impression. *Arch Phys.Med Rehabil.* 2006;87:897-903.

3. Benoist M. The natural history of lumbar degenerative spinal stenosis. *Joint Bone Spine.* 2002;69:450-7.

4. Kent DL, Haynor DR, Larson EB et al. Diagnosis of lumbar spinal stenosis in adults: a metaanalysis of the accuracy of CT, MR, and myelography. *AJR Am J Roentgenol.* 1992;158:1135-44.

5. Lurie JD, Birkmeyer NJ, Weinstein JN. Rates of advanced spinal imaging and spine surgery. *Spine.* 2003;28:616-20.

6. Deyo RA, Mirza SK. Trends and variations in the use of spine surgery. *Clin.Orthop.Relat Res.* 2006;443:139-46.:139-46.

7. Deyo RA, Ciol MA, Cherkin DC et al. Lumbar spinal fusion. A cohort study of complications, reoperations, and resource use in the Medicare population. *Spine.* 1993;18:1463-70.

8. Ciol MA, Deyo RA, Howell E et al. An assessment of surgery for spinal stenosis: time trends, geographic variations, complications, and reoperations. *J. Am Geriatr.Soc.* 1996;44:285-90.

9. Fanuele JC, Birkmeyer NJ, Abdu WA et al. The impact of spinal problems on the health status of patients: have we underestimated the effect? *Spine.* 2000;25:1509-14.

10. Taylor VM, Deyo RA, Cherkin DC et al. Low back pain hospitalization. Recent United States trends and regional variations. *Spine.* 1994;19:1207-12.

11. Turner SE, Eastwood PR, Cecins NM et al. Physiologic Responses to Incremental and Self-Paced Exercise in COPD; A Comparison of Three Tests. *Chest* 2004;766-73.

12. Deyo RA, Gray DT, Kreuter W et al. United States trends in lumbar fusion surgery for degenerative conditions. *Spine.* 2005;30:1441-5.

13. Whitman JM, Flynn TW, Fritz JM. Nonsurgical management of patients with lumbar spinal stenosis: a literature review and a case series of three patients managed with physical therapy. *Phys.Med Rehabil Clin.N.Am.* 2003;14:77-vii.

14. Atlas SJ, Delitto A. Spinal stenosis: surgical versus nonsurgical treatment. *Clin.Orthop.Relat Res.* 2006;443:198-207.:198-207.

15. Deen HG, Jr., Zimmerman RS, Lyons MK et al. Test-retest reproducibility of the exercise treadmill examination in lumbar spinal stenosis. *Mayo Clin.Proc.* 2000;75:1002-7.

16. Yamashita K, Hayashi J, Ohzono K et al. Correlation of patient satisfaction with symptom severity and walking ability after surgical treatment for degenerative lumbar spinal stenosis. *Spine.* 2003;28:2477-81.

17. Katz JN, Lipson SJ, Chang LC et al. Seven- to 10-year outcome of decompressive surgery for degenerative lumbar spinal stenosis. *Spine.* 1996;21:92-8.

18. Katz JN, Lipson SJ, Lew RA et al. Lumbar laminectomy alone or with instrumented or noninstrumented arthrodesis in degenerative lumbar spinal stenosis. Patient selection, costs, and surgical outcomes. *Spine.* 1997;22:1123-31.

19. Iversen MD, Katz JN. Examination findings and self-reported walking capacity in patients with lumbar spinal stenosis. *Phys.Ther.* 2001;81:1296-306.

20. Katz JN, Stucki G, Lipson SJ et al. Predictors of surgical outcome in degenerative lumbar spinal stenosis. *Spine.* 1999;24:2229-33.

21. Frazier DD, Lipson SJ, Fossel AH et al. Associations between spinal deformity and outcomes after decompression for spinal stenosis. *Spine.* 1997;22;2025-9.

22. Simotas AC. Nonoperative treatment for lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;153-61.

23. Hansraj KK, O'Leary PF, Cammisa FP, Jr. et al. Decompression, fusion, and instrumentation surgery for complex lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;18-25.

24. Hansraj KK, Cammisa FP, Jr., O'Leary PF et al. Decompressive surgery for typical lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;10-7.

25. Herno A, Airaksinen O, Saari T et al. Surgical results of lumbar spinal stenosis. A comparison of patients with or without previous back surgery. *Spine.* 1995;20:964-9.

26. Herno A, Airaksinen O, Saari T et al. Lumbar spinal stenosis: a matched-pair study of operated and non-operated patients. *Br.J.Neurosurg.* 1996;10:461-5.

27. McGregor AH, Hughes SP. The evaluation of the surgical management of nerve root compression in patients with low back pain: Part 1: the assessment of outcome. *Spine.* 2002;27:1465-70.

28. Herno A, Saari T, Suomalainen O et al. The degree of decompressive relief and its relation to clinical outcome in patients undergoing surgery for lumbar spinal stenosis. *Spine.* 1999;24:1010-4.

29. Shabat S, Leitner Y, Nyska M et al. Surgical treatment of lumbar spinal stenosis in patients aged 65 years and older. *Arch.Gerontol.Geriatr.* 2002;35:143-52.

30. Herno A, Airaksinen O, Saari T. Computed tomography after laminectomy for lumbar spinal stenosis. Patients' pain patterns, walking capacity, and subjective disability had no correlation with computed tomography findings. *Spine.* 1994;19:1975-8.

31. Herno A, Partanen K, Talaslahti T et al. Long-term clinical and magnetic resonance imaging follow-up assessment of patients with lumbar spinal stenosis after laminectomy. *Spine.* 1999;24:1533-7.

32. Tenhula J, Lenke LG, Bridwell KH et al. Prospective functional evaluation of the surgical treatment of neurogenic claudication in patients with lumbar spinal stenosis. *J.Spinal Disord.* 2000;13:276-82.

33. Airaksinen O, Herno A, Turunen V et al. Surgical outcome of 438 patients treated surgically for lumbar spinal stenosis. *Spine.* 1997;22:2278-82.

34. Deen HG, Jr., Zimmerman RS, Lyons MK et al. Measurement of exercise tolerance on the treadmill in patients with symptomatic lumbar spinal stenosis: a useful indicator of functional status and surgical outcome. *J Neurosurg.* 1995;83:27-30.

35. Barz T, Melloh M, Staub L et al. The diagnostic value of a treadmill test in predicting lumbar spinal stenosis. *Eur Spine J.* 2008;..

36. Fritz JM, Erhard RE, Delitto A et al. Preliminary results of the use of a two-stage treadmill test as a clinical diagnostic tool in the differential diagnosis of lumbar spinal stenosis. *J Spinal Disord.* 1997;10:410-6.

37. Adamova B, Vohanka S, Dusek L. Differential diagnostics in patients with mild lumbar spinal stenosis: the contributions and limits of various tests. *Eur Spine J.* 2003;12:190-6.

38. Whitehurst M, Brown LE, Eidelson SG et al. Functional mobility performance in an elderly population with lumbar spinal stenosis. *Arch Phys.Med Rehabil.* 2001;82:464-7.

39. Herno A, Airaksinen O, Saari T et al. Computed tomography findings 4 years after surgical management of lumbar spinal stenosis. No correlation with clinical outcome. *Spine.* 1999;24:2234-9.

40. Jensen OH, Schmidt-Olsen S. A new functional test in the diagnostic evaluation of neurogenic intermittent claudication. *Clin.Rheumatol.* 1989;8:363-7.

41. Pratt RK, Fairbank JC, Virr A. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. *Spine.* 2002;27:84-91.

42. Stucki G, Liang MH, Fossel AH et al. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J.Clin.Epidemiol.* 1995;48:1369-78.

43. Walsh TL, Hanscom B, Lurie JD et al. Is a condition-specific instrument for patients with low back pain/leg symptoms really necessary? The responsiveness of the Oswestry Disability Index, MODEMS, and the SF-36. *Spine.* 2003;28:607-15.

44. Gronblad M, Hupli M, Wennerstrand P et al. Intercorrelation and test-retest reliability of the Pain Disability Index (PDI) and the Oswestry Disability Questionnaire (ODQ) and their correlation with pain intensity in low back pain patients. *Clin.J.Pain.* 1993;9:189-95.

45. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine.* 2000;25:2940-52.

46. Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain.* 1996;65:71-6.

47. Gunzburg R, Keller TS, Szpalski M et al. Clinical and psychofunctional measures of conservative decompression surgery for lumbar spinal stenosis: a prospective cohort study. *Eur.Spine J.* 2003;12:197-204

CHAPTER 2

BACKGROUND

2.1 Lumbar Spinal Stenosis Definition and Clinical Presentation

Recently, the North American Spine Society (NASS) released clinical guidelines for the diagnosis and treatment of degenerative lumbar spinal stenosis (LSS).[1] These guidelines provided the following working definition of LSS:

> *Degenerative lumbar spinal stenosis is a clinical syndrome of buttock or lower extremity pain, which may occur with or without back pain, associated with diminished space available for the neural and vascular elements of the lumbar spine* [1] *(p.11)*

Degenerative spinal stenosis generally presents in individuals during their sixth or seventh decades of life (>50 years of age).[2;3] Usually LSS is a chronic condition, beginning with years of low back pain, interspersed with acute periods of intense pain, with or without lower extremity involvement.[4] It is common for symptoms of LSS to wax and wane, with acute exacerbations followed by a return to baseline symptoms.[3]

Pain is typically the primary symptom and main reason for patients with LSS seeking care.[5;5-7] Patients suffering from LSS typically report pain in the lower back, buttocks, thighs and legs.[8] The discomfort is often described as a cramping or burning sensation.[8] Patients with LSS also report physical impairments including poor balance, sensory loss (numbness/tingling) and muscle weakness in the buttocks and lower extremities.[6;9;10] Symptoms are generally intermittent and posture dependent, appearing with standing and lumbar extension, exacerbated by walking, and relieved by rest in a flexed or seated position.[4;8] Given the exacerbation of symptoms by walking, patients often have limited walking capacities.[6;10-12]

Symptom patterns vary among LSS patients,[4] from dull and aching pain in the sacroiliac and posterolateral thigh areas that is of gradual onset, to sharp radicular type pain in the lower extremities.[3;4;13] In central stenosis cases, the pain may be bilateral, although not symmetric, whereas patients with strictly foraminal involvement report symptoms closely resembling uni-lateral radiculopathy.[14;15]

Pain or discomfort can ascend or descend within the body and is localized poorly.[13]

2.1.1 *Neurogenic Claudication*

Neurogenic claudication is the cardinal manifestation and most specific symptom of LSS.[8;16;17] Neurogenic claudication consists of the progressive onset of radicular pain, paresthesias, numbness, weakness and tingling in the low back, buttocks and legs initiated by standing, walking or lumbar extension.[4;8;18] Most patients report symptomatic relief with forward flexion, sitting and/or recumbency.[1] Given the pain and discomfort associated with neurogenic claudication during walking, many LSS patients have limited walking capacities, require walking aids, or avoid walking altogether.

2.2 Diagnosis of LSS

There is no criterion standard for the clinical diagnosis of LSS.[19] In the absence of valid objective criteria it has been suggested that expert opinion be considered the 'gold standard' in LSS diagnosis, given that it provides a reasonable method of establishing a clinical diagnosis.[20] Diagnosis of the clinical syndrome of LSS is generally accomplished using a combination of clinical signs determined from examination of history, physical examination and imaging studies.[21]

2.2.1 *Clinical Historical and Physical Findings in LSS Patients*

Clinicians generally conduct a history and physical examination of potential LSS patients aimed at detecting findings characteristic of LSS. There are a number of historical and physical findings consistent with LSS which may lead to a diagnosis. Historical findings found to be most strongly associated with LSS are age (>50), severe lower extremity pain, absence of pain when seated, improvement of pain with sitting/flexion and worsened symptoms with walking.[1;21] Physical findings found to be most highly associated with LSS include: wide-based gait, abnormal Romberg test (balance), neuromuscular signs in the lower extremity including decreased strength (weakness), sensory deficits (numbness) and absent or decreased Achilles and patellar reflexes.[1;3;21] Thigh pain with 30 seconds of standing lumbar extension has also been shown to be strongly

associated with LSS.[21] Neurogenic claudication is the most specific symptom of LSS, but can only be observed when a patient is actually walking.[16] However, clinicians do not uniformly utilize observational tests of walking in LSS diagnosis. In addition to the aforementioned physical findings, some clinicians also employ electrodiagnostic tests in diagnosis of LSS. Electrodiagnostic methods, such as electromyography (EMG), are useful in testing the physiologic consequences of stenosis and in differential diagnosis from other disorders such as vascular claudication.[22]

2.2.2 *Imaging*

Definitive diagnostic information relating to LSS is most readily obtained from lumbar spine imaging.[4] The most appropriate non-invasive test for imaging LSS is magnetic resonance imaging (MRI).[1] MRI allows examination of the size, shape and anatomic relationships of spinal and neural elements.[4] Computed tomography (CT) is also commonly used in diagnosis of patients with LSS when MRI is contraindicated, or unavailable. Myelography has also been used extensively with LSS populations, however it is used less frequently given the technological advances of MRI and CT.[23]

Although imaging reports showing compression are a necessary component of LSS diagnosis, alone they are not sufficient. Spinal stenosis is a clinical condition, not a radiological finding or diagnosis. Up to 21% of subjects with stenosis demonstrated on MRI are asymptomatic.[24] In fact, most elderly individuals show some degree of spinal degeneration and stenosis on imaging, however, most are symptom free.[24] Therefore, it is necessary to use imaging studies in combination with an examination of history and clinical presentation, given that to date, a clear relationship has not be established between the severity of clinical symptoms and the degree of anatomical stenosis determined by imaging studies.[8;11;25-42] Results of imaging studies are often non-specific and cannot ensure that symptoms are arising from compression demonstrated on imaging.[24]

2.3 Patho-Anatomy of LSS

The adult spinal cord is housed in the central spinal canal, created by the vertebrae of the spine. The central lumbar spinal canal is bounded anteriorly by the lumbar vertebral bodies and intervertebral discs, anterolaterally by the pedicles, and posteriorally by the laminae of the vertebral arch, the ligamentum flavum and facet joints (zygoapophyseal joints).[8;43] The spinal cord typically ends around the level of the L1 vertebrae, and continues as the cauda equina, which is composed of multiple nerve roots. The nerve roots of the cauda equina exit the central spinal canal at each vertebral level, passing through the lateral recess zone to their specific intervertebral foramina.[8;44] These nerve roots are accompanied by small radicular arteries and veins. The lumbosacral nerve roots provide the neural control for the lower extremities.

LSS can be classified as congenital (developmental) and/or acquired.[44] Congenital stenosis is uncommon, and involves spinal canal narrowing caused by congenital abnormalities or a disorder in postnatal development.[45] The majority of LSS cases are acquired degenerative stenosis, resulting from degenerative changes of the aging spine.[8] Degenerative changes occur within the three joint complexes at levels L1 through S1 (Each three-joint complex is composed of two facet joints and an intervertebral disc).[8] Degenerative changes occurring in these three-joint complexes can lead to a narrowing of the central spinal canal, lateral recesses and/or intervertebral foramen causing compression of the associated neurovascular structures and resulting in central, lateral, and foraminal stenosis, respectively. Individuals can have central, lateral or foraminal stenosis in isolation, or in varying combinations. Combination stenosis is a term used to describe a combination of central and either lateral or foraminal stenosis.

It has been suggested that the degenerative process underlying stenosis often begins with changes in the intervertebral discs, moving subsequently to the facet joints.[17] Potentially stenotic changes associated with the intervertebral disc include loss of disc height and bulging. Other degenerative changes precipitating stenosis include the formation of osteophytes, as well as hypertrophy of facet joints, pedicles, laminae and ligamentum flavum.[16;44;46] In some cases, a defect in

the pedicles, called spondylolisis, may lead to spondylolisthesis, the anterior displacement of one vertebrae relative to another, which can further narrow an already diminished canal or foraminal space.[18] Any combination of these age related changes within the spine can lead to narrowing of the canal and intervertebral foramen, and subsequent compression of the neurovascular elements. Acquired degenerative stenosis may also occur post-surgically (excess scar tissue or proliferation of bone), as a result of infection or after trauma (fracture).[45]

It is important to note that although narrowing of the spinal canal and foramen is a necessary component of LSS, it alone is not sufficient for the disorder to be expressed.[8] In order for the LSS to be expressed symptomatically, the degree of narrowing must be such that there is enough compression of the neurovascular contents of the canal to elicit compromise in sensory and motor nerve function.[8]

2.4 Pathophysiology of Neurogenic Claudication

Neurogenic claudication is almost always associated with central or combination lumbar stenosis.[8] The exact physiological mechanism of neurogenic claudication in central lumbar stenosis is currently unknown. Two main theories have been proposed to explain the pathophysiology of neurogenic claudication: the ischemic theory and the venous stasis (stagnant hypoxia) theory.[8] These theories are both based to some degree on the mechanical compression of nervous and vascular structures in stenotic canals.

2.4.1 *The Ischemic Theory of Neurogenic Claudication*

The nerve roots of the cauda equina receive metabolic requirements from the cerebral spinal fluid (CSF) via diffusion from arteries located on the surface of the nerve roots themselves.[4] It has been suggested that this anatomic arrangement puts nerve roots at risk for ischemia in the presence of stenosis related extrinsic compression.[4] Indeed, studies have demonstrated that compression of cauda equina nerve roots at pressures similar to those observed in LSS can decrease blood flow, thus leading to ischemic conditions which slow nervous conduction.[47] It is suggested that this nerve root ischemia, owing to extrinsic compression of

microvasculature in stenosis, can cause the symptoms of neurogenic claudication, including paresthesias, pain and weakness.[48;49]

The ischemic theory is of special importance for explaining difficulty during walking, given that as the metabolic demands of the cauda equina increase, potential for ischemia incresases. The ischemic theory suggests that the metabolic demands of walking cannot be met by local nerve root vasculature because blood flow is compromised by compression.[8;23;50] This lack of blood flow and subsequent decreased delivery of substrates may cause ischemia within the lumbosacral nerve roots, leading to pain, sensory loss and motor deficiencies during walking.[48;50]

2.4.2 *The Venous Stasis or Stagnant Anoxia Theory of Neurogenic Claudication*

Studies using myeloscopy have demonstrated venous congestion, rather than ischemia, in patients reporting neurogenic claudication.[51] The venous stasis theory suggests that the mechanism underlying neurogenic claudication is inadequate oxygenation and accumulation of metabolites in the cauda equina due to venous pooling in *multi-level* stenosis.[18;52] It is thought that the structures causing mechanical compression of nerves may also compress the veins that exit the central canal with the nerve roots, causing venous pooling, entrapment of the cerebral spinal fluid and decreased venous return.[18;52] Porter et al.[18] provided evidence from a porcine model which showed that venous pooling between two levels of stenosis transitions to venous engorgement of the nerve roots during walking. This venous engorgement in turn prevents the expected arteriolar vaso-dilatory response to activity. This blunted vaso-dilatory response during walking leads to hypoxia, decreased metabolic exchange and decreased nutritional supply to the nerve roots, causing subsequent nerve conduction failure.[18;52]

2.4.3 *Posture and Neurogenic Claudication*

In many patients with LSS and neurogenic claudication, assuming a position of lumbar extension or lordosis is enough to provoke symptoms, which are alleviated by flexion.[17] It is suggested that symptoms of neurogenic claudication are exacerbated by transient mechanical compression of the cauda equina by degenerative intervertebral discs (anterior) and thickened ligamentum

flavum (posterior) when the lumbar spine is extended and lordosis is accentuated.[53] The direct compression of the cauda equina during extension is thought to inhibit function in the sensory and motor neurons of the lumbosacral nerve roots, precipitating symptoms of pain, numbness, tingling and weakness, which are relieved during flexion.[53]

2.5 Natural History of LSS

The natural history of untreated LSS is largely unknown, given that most people diagnosed with LSS seek some sort of treatment.[54] Very few studies have focused on the natural history of LSS. The recent North American Spine Society (NASS) clinical guidelines examined the limited literature in this area and concluded that the natural history in clinically mild to moderate LSS can be favourable in about one third to one half of patients.[1] It was also suggested that in mild to moderate stenosis, rapid or catastrophic neurological decline is rare.[1] However, the authors concluded that insufficient evidence exists to draw any conclusions regarding the natural history of clinically or radiologically severe stenosis. This is likely because most individuals with severe stenosis pursue surgical treatment, rendering the study of natural history implausible. Other reviews of natural history suggest that the condition may deteriorate in some patients, improve in up to one third, with the majority of patients remaining unchanged for up to 8 years follow up.[5;16;54-58] It was concluded in a recent report by Haig et al.[59] that clinical spinal stenosis is a fluctuating and potentially improving continuum, where current function of patients predicts future function. A number of studies focusing on the natural history of non-surgically treated LSS also examined changes in walking capacity over time. It appears that both the overall condition and walking capacity remain largely unchanged over time, for up to 10 years.[5]

2.6 Walking Capacity in LSS

As previously discussed, the primary symptom manifestation of degenerative LSS is neurogenic claudication. Neurogenic claudication causes pain, numbness and weakness in the lower back, buttocks and legs during walking, and often leads to reduced walking capacities.[7;60] Walking is a focus for

both LSS patients and clinicians, given the potential for decreased walking capacity to negatively affect patients' health related quality of life.[7] The ability of patients with LSS to walk for activities of daily living is not something that can be overlooked, especially given that self-reported walking difficulty was found to be the strongest and only independent predictor of patient satisfaction post-surgery for LSS.[61] Accurate measurement of walking capacity is important for assessment of LSS treatment outcomes and for monitoring the natural progression of the condition.

2.6.1 *Definition of the Construct of Walking Capacity in LSS*

For the purpose of this thesis research, *walking capacity* was defined as the distance a person with LSS is able to walk without support on a level surface at a self-selected speed before being forced to stop due to symptoms of LSS. This definition will be used throughout the thesis research to follow. Walking distance was selected rather than walking time because distance was considered to have more impact on function and activities of daily living.

2.6.2 *Measuring Walking Capacity for Outcomes of Treatment and Diagnosis*

Given that neurogenic claudication is the primary symptom of LSS, walking distance is commonly used as a measure of condition severity and outcome of treatment.[31;62-66] Clinicians must be able to accurately measure walking capacity in LSS patients in order to know whether treatment has been successful or whether the condition is changing over time. There has been a recent trend toward emphasizing walking capacity as a key outcome indicator for patients with LSS who are receiving both surgical and non-surgical treatments.[67] Most recent research reports focusing on treatment for LSS report some measure of walking capacity.[60;62;68-70] As such, there is a growing demand for valid and reproducible outcome-based measures of walking in LSS populations.[63;71] Given that radiological reports of stenosis (MRI/CT) do not correlate well with clinical outcomes,[11;25-42] including walking,[26;30;31;38;66;72] there is a need for valid measures of both. If clinical outcome or progression of the condition cannot be determined or monitored using imaging alone, this confirms the need for psychometrically

sound measures of the important clinical factors associated with LSS, including walking capacity.

Given the lack of an established relationship between imaging and clinical examination, certain aspects of walking, such as walking induced symptoms, are highly relevant components in diagnosis of LSS.[73] Thus, many clinicians assess walking capacity, either through patient self-reporting or observationally in the process of diagnosing this condition. In attempting to diagnose LSS, the difficulty often lies in the absence of clinical symptoms at rest, with limited function and pain only manifested during physical activity such as walking.[73] As such, it has been suggested that measurement of walking capacity observationally, such as with a treadmill test, will simulate the physical strain and limitations experienced in patients' day to day lives.[2;65;73] Observational measurement of walking is therefore useful in diagnosis given that it allows clinicians to assess individual walking capacity, while observing and recording clinical symptoms such as neurogenic claudication, which are specific to and characteristic of LSS.[2] Clinical diagnostic testing of walking capacity for potential LSS patients may include monitoring of the time to onset of symptoms with exercise, total distance and the nature of walking induced symptoms.[1] It has been suggested that observational measurement of walking using a treadmill test may be useful in the differential diagnosis of LSS, such as when differentiating between neurogenic and vascular types of claudication.[2;23;63;65] However, no guidelines or standards for how to use walking capacity in LSS diagnosis have been reported in the literature.

2.7 Treadmill Measurement of Walking Capacity in LSS

Treadmill testing of patients with symptomatic LSS has been shown to be safe; inexpensive, quantifiable and easily administered.[2;63;65;67] Treadmill tests have been used with LSS populations for diagnosis,[2;63;65;73;74] defining baseline clinical status [2;12;23;30;31;39;63;65-67;73-75] and for assessing outcomes of surgical and pharmacological treatment.[23;30;31;39;66;75] As previously mentioned, treadmills are especially useful in patients with LSS given that most symptoms, including neurogenic claudication, can only be observed during a physical activity such as

walking. Therefore many find it useful to observationally measure walking using a treadmill test during diagnosis and assessment of treatment outcomes.

However, there is not one treadmill protocol consistently reported in the LSS literature. Rather, protocols vary in speed, duration and incline (Table 2-1). Reported speeds range from 2km/h to 4.8km/h (1.2-3mph).[2;12;65;73;75] Maximum test duration ranges from 12 to 15 minutes, with the exception of the test by Whitehurst et al. who used 70% of maximum heart rate as the termination point.[12] Most protocols reported in the literature were conducted at 0% grade, with the exception of Whitehurst et al. [12] who increased grade by 1% per minute throughout the test. A brief review of treadmill protocols reported in the literature can be found in Table 2-1. As evidenced in the Table 2-1, most of the treadmill protocols have been used in evaluating outcomes of surgery for LSS.[30;31;63;65;66;75] A few studies have also examined the utility of treadmill tests in LSS diagnosis.[2;23;73;74]

2.7.1 *Treadmill Protocols for which there Exists Psychometric Evidence*

Although the use of treadmill protocols has been reported in the literature, the North American Spine Society (NASS) clinical guidelines for LSS concluded that there is insufficient scientific evidence to support the use of treadmill testing as an outcome measure for LSS.[1] A limited amount of research has been reported examining the psychometric properties of treadmill testing in LSS, including studies by Deen et al,[63;65;67] Tenhula et al,[75] and Barz et al.[73] In terms of validity, treadmill walking has yet to be compared to an observational test of walking on level ground in an LSS population. As such, we do not know if any of the reported protocols are valid for use with the LSS population. A review of the most relevant treadmill protocols, including only those for which psychometric analyses have been conducted, follows.

Deen et al.[63;65;67] have provided the most comprehensive reports of treadmill testing in LSS, including details regarding how the protocol was selected. These studies were conducted to determine whether treadmill testing was effective in evaluating baseline functional status and outcome of surgery for LSS.[63;65] The authors examined what parameters would best elicit neurogenic

claudication and a symptom limited end point in treadmill testing with LSS patients, keeping in mind the disabled and often elderly nature of LSS patients.

In order to select a protocol appropriate for LSS patients, Deen et al. conducted a series of preliminary studies with the Bruce treadmill protocol.[83] The Bruce protocol [83] has traditionally been used to assess patients with cardiopulmonary conditions. Deen et al.[63;65] found that the Bruce protocol was not well suited to persons with LSS, given that the starting speed (1.7 mph or 2.7km/h) and ramp incline during the test emphasized cardiopulmonary function over musculoskeletal performance, imposing unrealistic demands on LSS patients.[63] It was also shown that while performing the Bruce protocol, patients with LSS developed cardiovascular symptoms and stopped the walking test before they developed leg pain or signs of neurogenic claudication. Thus, the tests were thought to be limited by cardiovascular fitness and not symptoms of LSS. Therefore, Deen et al. adopted a lower speed and level walking surface for the elicitation of neurogenic claudication and facilitation of an LSS symptom limited test.[63;65]

The treadmill protocol chosen by Deen et al. included two walking trials, conducted at 0% grade.[63;65] The first trial was conducted at a speed of 2km/h (1.2 mph), and the second at a self-selected speed (speed ranged from 0.4 to 2.2mph). The protocol specified that patients walk with an upright posture and avoid holding the handrails. The authors recorded time to first symptoms, time to severe symptoms and the nature of symptoms (fatigue, back pain, leg pain). The test was stopped at the onset of severe symptoms or after 15 minutes. Severe symptoms were said to be "a level of discomfort that would make patients stop their activities in usual life situations".[63;65]

The end point of fifteen minutes was chosen because Deen et al. found that no new information was provided by lengthening the test past 15 minutes. In preliminary studies it was shown that patients who walked for longer periods of time (30 minutes) were unlikely to develop neurological symptoms with further walking, and that the ability to walk for 15 minutes after LSS surgery was a reasonable indication of good surgical results.[63;65] It was suggested that increasing

treadmill speed or ramp incline might elicit symptoms in patients who were asymptomatic after 15 minutes. This suggestion might be useful in examining function in LSS patients who are not undergoing surgery, and thus whose symptoms may not be as severe or might not be elicited in such a short period of time (<15 min). However, Fritz et al.[2] demonstrated that due to the increase in canal diameter occurring with forward flexion of the spine, an increase in incline may actually alleviate symptoms of LSS and thus increase walking time.

The results of Deen's studies revealed that findings were almost identical in the two trials, suggesting that only one trial is needed at either of the speeds (1.2mph or self-selected).[63] However, the use of a self-selected speed is likely more representative of normal walking conditions. Both studies concluded that walking assessment using the treadmill protocol is a safe, easily administered and quantifiable means of assessing baseline functional walking status and surgical outcome in patients with neurogenic claudication due to LSS. [63;65]

Deen et al.[67] also investigated the test re-test reproducibility of this treadmill protocol. In this prospective study, 28 patients with severe LSS underwent treadmill testing, first at a walking speed of 1.2 mph and then at the patient's preferred walking speed. All patients had a second treadmill examination or "re-test" over a period of 1 to 4 days. Time to first symptoms and total ambulation time were measured. Differences between the baseline examination and the re-test examination were assessed using the concordance correlation coefficient (CCC). The CCCs for total walking time and time to first symptoms were 0.89 and 0.98 for the 1.2mph trial, and 0.98 and 0.96 for the preferred walking speed trial.[67] The test was deemed to have a high degree of reproducibility. However, validity of this treadmill protocol has not been investigated.

The study by Tenhula et al.[75] was designed to assess outcome of surgery for LSS. The authors were also interested in the correlation between self-reported assessments of physical function, a 100 mm visual analog scale for pain severity (VAS) and the Oswestry Disability Index, with functional changes noted in the treadmill test. Over 5 years, 32 patients with LSS underwent functional and self-

report assessments before and after surgery. Patients completed a treadmill test, the pain VAS and the Oswestry Disability Index forms pre and post-operatively. The following treadmill protocol was conducted at 0% grade: 10 minutes at 3.6km/h (2 mph), 5 minutes at 4.1km/h (2.5 mph), and 5 minutes at 5km/h (3mph). The time to onset or increase in symptoms, total time and total distance were recorded, as well as the location, severity and type of symptom (pain, parenthesis, weakness) on a body chart and VAS. The reason for stopping the test was recorded if the patient was unable to walk for the duration of the 20-minute test.

In terms of validity, a strong correlation (r=0.88) was found between total time on the treadmill test and post-operative self-reported pain on the VAS.[75] It was also found that patients' post-operative Oswestry scores correlated significantly with improvement in treadmill test parameters, although a value for the correlation coefficient was not reported. Tenhula et al.[75] suggested that the treadmill test may be used as an objective measure of post-operative outcome for surgery for LSS. They also suggested that functional testing using a treadmill may become a tool in the decision making process and evaluation for surgical treatment of LSS. Reproducibility was not investigated with this protocol. The NASS guidelines concluded that this study provided fair evidence supporting treadmill testing in the assessment of walking capacity as a functional measure of surgical outcome.[1] (p. 46)

2.7.2 *Examining the Diagnostic Value of Treadmill Testing in LSS*

Barz et al.[73] examined the relationship between total distance walked during a treadmill test and dural cross sectional area (measured on MRI) in 25 patients undergoing decompressive surgery for LSS. The protocol for this study consisted of treadmill walking at 1.1mph (1.8km/h). The correlations between walking distance and the ODI, as well as self-reported walking capacity were also reported. In terms of validity, the treadmill walking distance correlated significantly ($p<0.05$) with the dural cross sectional area (r=0.53), the ODI (r=0.51), and self-reported walking capacity (distance) (r=0.62) ($p<0.05$). Patients

tended to overestimate their actual walking distance as judged by treadmill walking by a factor of three.[73]

2.7.3 *Other Observational Measures of Walking Capacity in LSS*

Other than treadmill protocols, the Shuttle Walking Test is the only observational measure of walking in LSS that has received any psychometric analysis.[84] The Shuttle Walking Test was developed to assess exercise tolerance in patients with chronic obstructive pulmonary disease and cardiac disorders.[85-87] Although this test has been used as an outcome measure in clinical trials involving patients with low back pain, it has yet to be used to evaluate walking in an LSS population. One study by Pratt et al.[84] examined the reliability of the Shuttle Walking Test over an interval of one week in a group of 17 patients with LSS, and found the test re-test reliability to be high (ICC=0.92). Although this test was found to be reliable in an LSS sample, it may not be appropriate for measuring walking in LSS, given that the nature of the Shuttle Walking Test is incremental, with speed increasing throughout the test. Much like the aforementioned Bruce protocol examined by Deen et al.,[63] the Shuttle Walking Test is intended as an exercise stress test to assess cardiovascular fitness, and not as a symptom limited test to assess the effect of neurogenic claudication on walking capacity in LSS patients.

2.8 Self-report Measures of Walking Capacity in LSS

Although observational testing of walking capacity in LSS is likely more accurate, self-report instruments remain prevalent in clinical practice and research. There are a number of self-report measures which have been used specifically to assess walking capacity in patients with LSS. The instruments used most frequently for this purpose include the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire,[88] the Oswestry Disability Index [89] and the Oxford Claudication Score.[84] In addition to these instruments, a number of other self-report measures have been used in research, including the Health Utilities Index Single Attribute Utility Score for Ambulation,[90] Yamashita's 100mm visual analog scale (VAS) assessing subjective difficulty with walking [71] and a number of single item questions addressing the construct of walking capacity.

According to the recent NASS clinical guidelines for LSS,[1] the measures deemed to be most appropriate for assessing outcomes of treatment for LSS are the Swiss Spinal Stenosis Questionnaire and the Oswestry Disability Index. However, on a scale which included the grades of A (good evidence), B (fair evidence), C (poor quality evidence) and I (insufficient evidence), both the Swiss Spinal Stenosis Questionnaire and the ODI were given a grade of B. This implies that although these instruments were recommended as most appropriate, further research is warranted focusing on the psychometric properties of these questionnaires.

2.8.1 *Physical Function Scale of the Swiss Spinal Stenosis Questionnaire*

The Swiss Spinal Stenosis Questionnaire was developed specifically for use with LSS populations.[88;91] It was designed to complement existing generic measures of lumbar spine disability and health status in the evaluation of patients with LSS.[88] It was suggested in the NASS guidelines that the Swiss Spinal Stenosis Questionnaire is currently the best and most specific outcome measure for use with LSS populations.[1] It was also concluded in the NASS guidelines that in future studies focusing on specific outcome measures for the treatment of LSS, the Swiss Spinal Stenosis Questionnaire could be a potential gold standard.[1]

The Swiss Spinal Stenosis Questionnaire was developed using judgmental analysis, based on a literature review and consensus of a panel of six LSS experts,[88] including four rheumatologists, a spine specialist orthopaedic surgeon and a behavioural scientist. The three scales used in the Swiss Spinal Stenosis Questionnaire address symptom severity, satisfaction, and physical function. The Physical Function Scale was designed specifically to evaluate physical function and walking in LSS patients.[88] The items in the Physical Function Scale assess distance walked and pain limited activities of daily living involving walking. The Physical Function Scale has been used often to assess walking in studies examining the effectiveness of interventions for LSS.[6;92-101]

The psychometric properties of the Physical Function Scale have received some attention in the literature (Table 4-1). Stucki et al. conducted the two primary studies investigating the validity and reproducibility of this scale.[88;91]

Subjects for both studies were recruited from an ongoing prospective multi-centre observational study of patients undergoing surgery for LSS. These studies found the Physical Function Scale to be reproducible and responsive to clinical change.[88;91] However, the only pieces of validity evidence currently available to support the use of the Physical Function Scale in the assessment of walking capacity in LSS include the correlations with physicians' assessment of walking capacity (r=0.47) and the physical dimension of the Sickness Impact Profile (r=0.49) reported in the study by Stucki et al.[88] No studies to date have been designed to specifically assess the construct validity of the Physical Function Scale for the measurement of walking capacity in LSS.

In addition to the studies by Stucki et al.,[88;91] Pratt et al.[84] examined the internal consistency and test re-test reliability of the Physical Function Scale in a group of 29 patients with LSS and neurogenic claudication over a test interval of one week. It was concluded that the scale possessed high test re-test reliability (ICC=0.82) and was internally consistent (Cronbach's α=0.87-0.89). In addition, Thornes et al.[102] recently examined the psychometric properties of the Physical Function Scale to provide evidence supporting a cross-cultural adaptation of the questionnaire for use in Norwegian research on LSS. Subjects for this study were 75 individuals with confirmed LSS who were referred for surgery in Norway.[102] All questionnaires were completed prior to surgery. In this study, the Physical Function Scale was correlated significantly with both the ODI (r=0.70) and the visual analog scale (VAS) for leg pain (r=0.41), as hypothesized, providing construct validity evidence for the use of the Physical Function Scale in assessing pain related disability.[102] The Physical Function Scale was also found to be reproducible, with an ICC of 0.89 (95% confidence interval, 0.79-0.95) over a test re-test interval of one week.[102] In summary, there is evidence to suggest that the Physical Function Scale is internally consistent and reproducible in LSS populations. However, limited evidence exists regarding the validity of the Physical Function Scale for the measurement of walking capacity in LSS.

2.8.2 *Oswestry Disability Index*

The Oswestry Disability Index (ODI) is a low back specific measure of the extent to which function is influenced by back pain.[89] This instrument is psychometrically sound, and is recommended as a standard measurement for assessing back pain related function.[1;89;103-105] It includes dimensions concerning pain and the effect of pain on functional activities, including personal care, lifting, sitting, standing, sleeping, social life, traveling and walking. Questions are rated using 5-point or 6-point Likert scales with higher scores representing greater pain or dysfunction. The questionnaire also contains one item which addresses the construct of pain limited walking capacity (distance).

The ODI has been used in many studies investigating overall function in patients with LSS.[26;31;66;75;106;106-110] It has been used once to evaluate walking distance specifically, by isolating the walking specific item in a study examining outcomes of surgery for LSS.[106] In terms of psychometric properties, the ODI has been validated for use in evaluating function in low back pain patients.[105;107;111-113] It has also been shown to be reliable in a group of patients with neurogenic claudication and LSS (ICC=0.89).[84] However, limited evidence exists regarding the use of the ODI for the evaluation of walking in patients with LSS. The only evidence available comes from two different studies examining the relationship between treadmill walking and the ODI and patients undergoing surgery for LSS.[73;75] Barz et al. reported a correlation between the ODI and total distance walked during the treadmill test (r=0.51),[73] while Tenhula reported a significant correlation between post-operative ODI scores and treadmill test parameters (correlation coefficients not reported).[75] Given the widespread use of the ODI in LSS and low back pain research, it was useful to determine if the ODI is valid for evaluating walking capacity, to allow for comparison between LSS studies and between different low back conditions.

2.8.3 *Health Utilities Index Mark 3 Ambulation Score*

The Health Utilities Index Mark 3 (HUI3) is a generic 15-item general health questionnaire based on a multi-attribute, preference-based system of assessing health status.[90] From the HUI3 one can calculate eight single attribute

scores: vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain. Each attribute consists of 5 to 6 levels, and can be transformed into single attribute utility scores, on a scale from 0-1 with 0 representing most disabled and 1 representing no disability. The single attribute utility score for ambulation is based on item 9 of the questionnaire.[90] This ambulation score has never been reported in literature related to LSS, yet was thought to tap the construct of walking capacity in LSS.

2.8.4 *Visual Analog Scale for Walking Difficulty*

Yamashita et al.[71] used a 100mm visual analog scale (VAS) to examine subjective walking difficulty in a study correlating patient satisfaction with self-reported walking difficulty post-surgery for LSS. Self-reported walking difficulty was found to be the strongest and only independent correlate of patient satisfaction (Spearman r=0.43) when the variables of post-operative back pain, leg pain, numbness and walking time were also considered.[71] This item has only been used in the original study and has not been psychometrically evaluated.

2.9 Summary

As previously mentioned, there is no 'gold standard' measure of walking capacity presently used with LSS patients. Further, the validity and reproducibility of the currently employed measures have not been adequately evaluated. As such, it was important to examine the psychometric properties of the instruments and measures currently used for this purpose. It appears that the measures most commonly used and supported by the clinical and research communities include the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire, the Oswestry Disability Index and treadmill tests of walking. Limited validity evidence exists regarding the use of any of these measures for the assessment of walking in patients with LSS. It is likely that these instruments and tests will continue to be used in LSS clinical practice and research, given the desire for researchers and clinicians to replicate previous measurement methods and for the sake of effective comparison. However, none of the aforementioned measures has actually been compared to an observational measure of walking capacity conducted in a realistic walking environment on level ground. It is likely

that a test where patients walk on level ground and at their own pace until forced to stop due to symptoms of LSS would be a better gold standard or criterion than treadmill walking for measuring walking capacity in this population. Such a test would be representative of walking in an everyday context, which is the issue at hand when assessing patients' walking capacity. It is apparent that investigation into the validity of the measures currently used to assess walking capacity in LSS populations was warranted.

Table 2-1. Treadmill Protocols used to Measure Walking in LSS

Author	Study Purpose	Speed	Duration	Incline	Variables measured
Deen et al. [76;77]	Outcome of surgery	2km/h (1.2 mph) (Additional trial with speed selected by subjects)	15 minutes†	0%	Time to onset of symptoms* Total time
Herno et al. [78]	Outcome of surgery	3.6 km/h (2.2 mph)	15 minutes	0%	Total distance
Tenhula et al. [79]	Outcome of surgery	-10 min @ 3.2 km/h (2mph) -5 min @ 4.0 km/h (2.5 mph) -5 min @ 4.8 km/h (3mph)	15 minutes	0%	Time to onset of symptoms Total time Total distance
Whitehurst et al. [80]	Comparing functional mobility between patients with LSS and healthy older adults	53.6 m/min (2mph)	Until 70% of maximum heart rate attained, or until subjects requested to stop	Increase of 1% per minute for the duration of the test	Total time
Murakami et al [81]	Effects of lipoprostaglandin E1 treatment for neurogenic claudication	Self-selected	Not specified	0%	Distance to onset of symptoms Total distance
Adamova et al. [82]	Diagnosis of mild LSS	-3 min @ 1.6 km/h (1.0 mph) -3 min @ 2.4 km/h (1.5mph) -3 min @ 3.2 km/h (2 mph) -3 min @ 4.0 km/h (2.5 mph)	12 minutes	0%	Total distance Total Time
Fritz et al. [2]	Diagnosis based on postural dependency of neurogenic claudication symptoms	Self-selected	2 trials of 10 minutes	First trial: 0% Incline trial: 15%	Time to onset of symptoms Total time Recovery time
Jensen et al. [74]	Diagnosis based on postural dependency of neurogenic claudication symptoms	1.8km/h	15 minutes	10% decline	Total distance
Barz et al.[73]	Comparison of treadmill distance and MRI in diagnosis of LSS	1.8km/h (1.1mph)	Not specified	0%	Total distance

* The nature and location of symptoms was recorded in all protocols, both before and after treadmill testing.
† Duration represents the maximum total time; All tests were symptom limited, meaning the test is terminated when the patient indicates that their symptoms are too severe continue.

2.10 References

1. Watters WC, III, Baisden J, Gilbert TJ et al. Degenerative lumbar spinal stenosis: an evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis. *Spine J.* 2008;8:305-10.

2. Fritz JM, Erhard RE, Delitto A et al. Preliminary results of the use of a two-stage treadmill test as a clinical diagnostic tool in the differential diagnosis of lumbar spinal stenosis. *J Spinal Disord.* 1997;10:410-6.

3. Thomas SA. Spinal stenosis: history and physical examination. *Phys.Med Rehabil Clin.N.Am.* 2003;14:29-39.

4. Binder DK, Schmidt MH, Weinstein PR. Lumbar spinal stenosis. *Semin.Neurol.* 2002;22:157-66.

5. Amundsen T, Weber H, Nordal HJ et al. Lumbar spinal stenosis: conservative or surgical management?: A prospective 10-year study. *Spine.* 2000;25:1424-35.

6. Iversen MD, Katz JN. Examination findings and self-reported walking capacity in patients with lumbar spinal stenosis. *Phys.Ther.* 2001;81:1296-306.

7. Lin SI, Lin RM. Disability and walking capacity in patients with lumbar spinal stenosis: association with sensorimotor function, balance, and functional performance. *J Orthop.Sports Phys.Ther.* 2005;35:220-6.

8. Chad DA. Lumbar spinal stenosis. *Neurol.Clin.* 2007;25:407-18.

9. Johnsson KE, Rosen I, Uden A. Neurophysiologic investigation of patients with spinal stenosis. *Spine.* 1987;12:483-7.

10. Stucki G, Liang MH, Lipson SJ et al. Contribution of neuromuscular impairment to physical functional status in patients with lumbar spinal stenosis. *J Rheumatol.* 1994;21:1338-43.

11. Jonsson B, Annertz M, Sjoberg C et al. A prospective and consecutive study of surgically treated lumbar spinal stenosis. Part I: Clinical features related to radiographic findings. *Spine.* 1997;22:2932-7.

12. Whitehurst M, Brown LE, Eidelson SG et al. Functional mobility performance in an elderly population with lumbar spinal stenosis. *Arch Phys.Med Rehabil.* 2001;82:464-7.

13. Hall S, Bartleson JD, Onofrio BM et al. Lumbar spinal stenosis. Clinical features, diagnostic procedures, and results of surgical treatment in 68 patients. *Ann.Intern.Med.* 1985;103:271-5.

14. Lee CK, Rauschning W, Glenn W. Lateral lumbar spinal canal stenosis: classification, pathologic anatomy and surgical decompression. *Spine.* 1988;13:313-20.

15. Reynolds AF, Weinstein PR, Wachter RD. Lumbar monoradiculopathy due to unilateral facet hypertrophy. *Neurosurgery.* 1982;10:480-6.

16. Atlas SJ, Delitto A. Spinal stenosis: surgical versus nonsurgical treatment. *Clin.Orthop.Relat Res.* 2006;443:198-207.:198-207.

17. Arbit E, Pannullo S. Lumbar stenosis: a clinical review. *Clin.Orthop.Relat Res.* 2001;137-43.

18. Porter RW. Spinal stenosis and neurogenic claudication. *Spine.* 1996;21:2046-52.

19. Haig AJ, Tong HC, Yamakawa KS et al. Spinal stenosis, back pain, or no symptoms at all? A masked study comparing radiologic and electrodiagnostic diagnoses to the clinical impression. *Arch Phys.Med Rehabil.* 2006;87:897-903.

20. Katz JN, Dalgas M, Stucki G et al. Diagnosis of lumbar spinal stenosis. *Rheum.Dis.Clin.North Am.* 1994;20:471-83.

21. Katz JN, Dalgas M, Stucki G et al. Degenerative lumbar spinal stenosis. Diagnostic value of the history and physical examination. *Arthritis Rheum.* 1995;38:1236-41.

22. Haig AJ, Tong HC, Yamakawa KS et al. The sensitivity and specificity of electrodiagnostic testing for the clinical syndrome of lumbar spinal stenosis. *Spine.* 2005;30:2667-76.

23. Adamova B, Vohanka S, Dusek L. Differential diagnostics in patients with mild lumbar spinal stenosis: the contributions and limits of various tests. *Eur Spine J.* 2003;12:190-6.

24. Boden SD, Davis DO, Dina TS et al. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg.Am.* 1990;72:403-8.

25. Jalovaara P, Lahde S, Iikko E et al. The significance of residual stenosis after decompression for lumbar spinal stenosis. *Ann.Chir Gynaecol.* 1989;78:304-8.

26. Herno A, Airaksinen O, Saari T. Computed tomography after laminectomy for lumbar spinal stenosis. Patients' pain patterns, walking capacity, and subjective disability had no correlation with computed tomography findings. *Spine.* 1994;19:1975-8.

27. Postacchini F. Management of lumbar spinal stenosis. *J Bone Joint Surg.Br.* 1996;78:154-64.

28. Bernard TN, Jr. Repeat lumbar spine surgery. Factors influencing outcome. *Spine.* 1993;18:2196-200.

29. Amundsen T, Weber H, Lilleas F et al. Lumbar spinal stenosis. Clinical and radiologic features. *Spine.* 1995;20:1178-86.

30. Herno A, Airaksinen O, Saari T et al. Computed tomography findings 4 years after surgical management of lumbar spinal stenosis. No correlation with clinical outcome. *Spine.* 1999;24:2234-9.

31. Herno A, Saari T, Suomalainen O et al. The degree of decompressive relief and its relation to clinical outcome in patients undergoing surgery for lumbar spinal stenosis. *Spine.* 1999;24:1010-4.

32. Ciol MA, Deyo RA, Howell E et al. An assessment of surgery for spinal stenosis: time trends, geographic variations, complications, and reoperations. *J Am Geriatr.Soc.* 1996;44:285-90.

33. Frazier DD, Lipson SJ, Fossel AH et al. Associations between spinal deformity and outcomes after decompression for spinal stenosis. *Spine.* 1997;22:2025-9.

34. Hu RW, Jaglal S, Axcell T et al. A population-based study of reoperations after back surgery. *Spine.* 1997;22:2265-70.

35. Herron LD, Mangelsdorf C. Lumbar spinal stenosis: results of surgical treatment. *J Spinal Disord.* 1991;4:26-33.

36. Lehmann TR, Spratt KF, Tozzi JE et al. Long-term follow-up of lower lumbar fusion patients. *Spine.* 1987;12:97-104.

37. Resnick DK, Choudhri TF, Dailey AT et al. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 5: correlation between radiographic and functional outcome. *J Neurosurg.Spine.* 2005;2:658-61.

38. Lohman CM, Tallroth K, Kettunen JA et al. Comparison of radiologic signs and clinical symptoms of spinal stenosis. *Spine.* 2006;31:1834-40.

39. Moon ES, Kim HS, Park JO et al. Comparison of the predictive value of myelography, computed tomography and MRI on the treadmill test in lumbar spinal stenosis. *Yonsei Med J.* 2005;46:806-11.

40. Geisser ME, Haig AJ, Tong HC et al. Spinal canal size and clinical symptoms among persons diagnosed with lumbar spinal stenosis. *Clin.J Pain.* 2007;23:780-5.

41. Uden A, Johnsson KE, Jonsson K et al. Myelography in the elderly and the diagnosis of spinal stenosis. *Spine.* 1985;10:171-4.

42. Weir B, de LR. Lumbar stenosis: analysis of factors affecting outcome in 81 surgical cases. *Can.J Neurol.Sci.* 1981;8:295-8.

43. Deyo RA, Weinstein JN. Low back pain. *N.Engl.J Med.* 2001;344:363-70.

44. Botwin K, Brown LA, Fishman M et al. Fluoroscopically guided caudal epidural steroid injections in degenerative lumbar spine stenosis. *Pain Physician.* 2007;10:547-58.

45. Ciricillo SF, Weinstein PR. Lumbar spinal stenosis. *West J Med.* 1993;158:171-7.

46. Tenhula J, Lenke LG, Bridwell KH et al. Prospective functional evaluation of the surgical treatment of neurogenic claudication in patients with lumbar spinal stenosis. *J Spinal Disord.* 2000;13:276-82.

47. Rydevik B. Neurophysiology of cauda equina compression. *Acta Orthop.Scand.Suppl.* 1993;251:52-5.:52-5.

48. Watanabe R, Parke WW. Vascular and neural pathology of lumbosacral spinal stenosis. *J Neurosurg.* 1986;64:64-70.

49. Evan JG. Neurogenic Intermittment Claudication. *Br Med J.* 1964;2:985-7.

50. Blau JN, Logue V. The natural history of intermittent claudication of the cauda equina. A long term follow-up study. *Brain.* 1978;101:211-22.

51. Ooi Y, Mita F, Satoh Y. Myeloscopic study on lumbar spinal canal stenosis with special reference to intermittent claudication. *Spine.* 1990;15:544-9.

52. Porter RW, Ward D. Cauda equina dysfunction. The significance of two-level pathology. *Spine.* 1992;17:9-15.

53. Alvarez JA, Hardy RH, Jr. Lumbar spine stenosis: a common cause of back and leg pain. *Am Fam.Physician.* 1998;57:1825-40.

54. Benoist M. The natural history of lumbar degenerative spinal stenosis. *Joint Bone Spine.* 2002;69:450-7.

55. Johnsson KE, Rosen I, Uden A. The natural course of lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 1992;82-6.

56. Johnsson KE, Uden A, Rosen I. The effect of decompression on the natural course of spinal stenosis. A comparison of surgically treated and untreated patients. *Spine.* 1991;16:615-9.

57. Mariconda M, Fava R, Gatto A et al. Unilateral laminectomy for bilateral decompression of lumbar spinal stenosis: a prospective comparative study with conservatively treated patients. *J Spinal Disord.Tech.* 2002;15:39-46.

58. Hurri H, Slatis P, Soini J et al. Lumbar spinal stenosis: assessment of long-term outcome 12 years after operative and conservative treatment. *J Spinal Disord.* 1998;11:110-5.

59. Haig AJ, Tong HC, Yamakawa KS et al. Predictors of pain and function in persons with spinal stenosis, low back pain, and no back pain. *Spine.* 2006;31:2950-7.

60. Shabat S, Arinzon Z, Folman Y et al. Long-term outcome of decompressive surgery for lumbar spinal stenosis in octogenarians. *Eur Spine J.* 2008;17:193-8.

61. Yamashita K, Ohzono K, Hiroshima K. Patient satisfaction as an outcome measure after surgical treatment for lumbar spinal stenosis: testing the validity and discriminative ability in terms of symptoms and functional status. *Spine.* 2006;31:2602-8.

62. Brussee P, Hauth J, Donk RD et al. Self-rated evaluation of outcome of the implantation of interspinous process distraction (X-Stop) for neurogenic claudication. *Eur Spine J.* 2008;17:200-3.

63. Deen HG, Jr., Zimmerman RS, Lyons MK et al. Measurement of exercise tolerance on the treadmill in patients with symptomatic lumbar spinal stenosis: a useful indicator of functional status and surgical outcome. *J Neurosurg.* 1995;83:27-30.

64. Jonsson B, Stromqvist B. Decompression for lateral lumbar spinal stenosis. Results and impact on sick leave and working conditions. *Spine.* 1994;19:2381-6.

65. Deen HG, Zimmerman RS, Lyons MK et al. Use of the exercise treadmill to measure baseline functional status and surgical outcome in patients with severe lumbar spinal stenosis. *Spine.* 1998;23:244-8.

66. Herno A, Partanen K, Talaslahti T et al. Long-term clinical and magnetic resonance imaging follow-up assessment of patients with lumbar spinal stenosis after laminectomy. *Spine.* 1999;24:1533-7.

67. Deen HG, Jr., Zimmerman RS, Lyons MK et al. Test-retest reproducibility of the exercise treadmill examination in lumbar spinal stenosis. *Mayo Clin.Proc.* 2000;75:1002-7.

68. Atlas SJ, Keller RB, Wu YA et al. Long-term outcomes of surgical and nonsurgical management of lumbar spinal stenosis: 8 to 10 year results from the maine lumbar spine study. *Spine.* 2005;30:936-43.

69. Fu YS, Zeng BF, Xu JG. Long-term outcomes of two different decompressive techniques for lumbar spinal stenosis. *Spine.* 2008;33:514-8.

70. Willen J, Wessberg PJ, Danielsson B. Surgical results in hidden lumbar spinal stenosis detected by axial loaded computed tomography and magnetic resonance imaging: an outcome study. *Spine.* 2008;33:E109-E115.

71. Yamashita K, Hayashi J, Ohzono K et al. Correlation of patient satisfaction with symptom severity and walking ability after surgical treatment for degenerative lumbar spinal stenosis. *Spine.* 2003;28:2477-81.

72. Goh KJ, Khalifa W, Anslow P et al. The clinical syndrome associated with lumbar spinal stenosis. *Eur Neurol.* 2004;52:242-9.

73. Barz T, Melloh M, Staub L et al. The diagnostic value of a treadmill test in predicting lumbar spinal stenosis. *Eur Spine J.* 2008;..

74. Jensen OH, Schmidt-Olsen S. A new functional test in the diagnostic evaluation of neurogenic intermittent claudication. *Clin.Rheumatol.* 1989;8:363-7.

75. Tenhula J, Lenke LG, Bridwell KH et al. Prospective functional evaluation of the surgical treatment of neurogenic claudication in patients with lumbar spinal stenosis. *J.Spinal Disord.* 2000;13:276-82.

76. Deen HG, Jr., Zimmerman RS, Lyons MK et al. Measurement of exercise tolerance on the treadmill in patients with symptomatic lumbar spinal stenosis: a useful indicator of functional status and surgical outcome. *J.Neurosurg.* 1995;83:27-30.

77. Deen HG, Zimmerman RS, Lyons MK et al. Use of the exercise treadmill to measure baseline functional status and surgical outcome in patients with severe lumbar spinal stenosis. *Spine* 1998;23:244-8.

78. Herno A, Saari T, Suomalainen O et al. The degree of decompressive relief and its relation to clinical outcome in patients undergoing surgery for lumbar spinal stenosis. *Spine* 1999;24:1010-4.

79. Hurri H, Slatis P, Soini J et al. Lumbar spinal stenosis: assessment of long-term outcome 12 years after operative and conservative treatment. *J.Spinal Disord.* 1998;11:110-5.

80. Whitehurst M, Brown LE, Eidelson SG et al. Functional mobility performance in an elderly population with lumbar spinal stenosis. *Arch.Phys.Med.Rehabil.* 2001;82:464-7.

81. Murakami M, Takahashi K, Sekikawa T et al. Effects of intravenous lipoprostaglandin E1 on neurogenic intermittent claudication. *J Spinal Disord.* 1997;10:499-504.

82. Adamova B, Vohanka S, Dusek L. Differential diagnostics in patients with mild lumbar spinal stenosis: the contributions and limits of various tests. *Eur.Spine J.* 2003;12:190-6.

83. Bruce RA, Kusumi F, Hosmer D. Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. *Am Heart J.* 1973;85:546-62.

84. Pratt RK, Fairbank JC, Virr A. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. *Spine.* 2002;27:84-91.

85. Keell SD, Chambers JS, Francis DP et al. Shuttle-walk test to assess chronic heart failure. *Lancet.* 1998;352:705.

86. Payne GE, Skehan JD. Shuttle walking test: a new approach for evaluating patients with pacemakers. *Heart.* 1996;75:414-8.

87. Singh SJ, Morgan MD, Scott S et al. Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax.* 1992;47:1019-24.

88. Stucki G, Daltroy L, Liang MH et al. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine.* 1996;21:796-803.

89. Fairbank JC, Couper J, Davies JB et al. The Oswestry low back pain disability questionnaire. *Physiotherapy.* 1980;66:271-3.

90. Feeny D, Furlong W, Boyle M et al. Multi-attribute health status classification systems. Health Utilities Index. *Pharmacoeconomics.* 1995;7:490-502.

91. Stucki G, Liang MH, Fossel AH et al. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J.Clin.Epidemiol.* 1995;48:1369-78.

92. Katz JN, Lipson SJ, Chang LC et al. Seven- to 10-year outcome of decompressive surgery for degenerative lumbar spinal stenosis. *Spine.* 1996;21:92-8.

93. Katz JN, Lipson SJ, Lew RA et al. Lumbar laminectomy alone or with instrumented or noninstrumented arthrodesis in degenerative lumbar spinal stenosis. Patient selection, costs, and surgical outcomes. *Spine.* 1997;22:1123-31.

94. Katz JN, Stucki G, Lipson SJ et al. Predictors of surgical outcome in degenerative lumbar spinal stenosis. *Spine.* 1999;24:2229-33.

95. Frazier DD, Lipson SJ, Fossel AH et al. Associations between spinal deformity and outcomes after decompression for spinal stenosis. *Spine.* 1997;22:2025-9.

96. Simotas AC. Nonoperative treatment for lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;153-61.

97. Hansraj KK, O'Leary PF, Cammisa FP, Jr. et al. Decompression, fusion, and instrumentation surgery for complex lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;18-25.

98. Hansraj KK, Cammisa FP, Jr., O'Leary PF et al. Decompressive surgery for typical lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;10-7.

99. Lyle MA, Manes S, McGuinness M et al. Relationship of physical examination findings and self-reported symptom severity and physical function in patients with degenerative lumbar conditions. *Phys.Ther.* 2005;85:120-33.

100. Zucherman JF, Hsu KY, Hartjen CA et al. A prospective randomized multi-center study for the treatment of lumbar spinal stenosis with the X STOP interspinous implant: 1-year results. *Eur Spine J.* 2004;13:22-31.

101. Zucherman JF, Hsu KY, Hartjen CA et al. A multicenter, prospective, randomized trial evaluating the X STOP interspinous process decompression system for the treatment of neurogenic intermittent claudication: two-year follow-up results. *Spine.* 2005;30:1351-8.

102. Thornes E, Grotle M. Cross-cultural adaptation of the Norwegian version of the spinal stenosis measure. *Eur Spine J.* 2008;..

103. Fairbank J. Use of Oswestry Disability Index (ODI). *Spine.* 1995;20:1535-7.

104. Fairbank J. Revised Oswestry Disability questionnaire. *Spine.* 2000;25:2552.

105. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine.* 2000;25:2940-52.

106. Shabat S, Leitner Y, Nyska M et al. Surgical treatment of lumbar spinal stenosis in patients aged 65 years and older. *Arch.Gerontol.Geriatr.* 2002;35:143-52.

107. McGregor AH, Hughes SP. The evaluation of the surgical management of nerve root compression in patients with low back pain: Part 1: the assessment of outcome. *Spine.* 2002;27:1465-70.

108. Herno A, Airaksinen O, Saari T et al. Surgical results of lumbar spinal stenosis. A comparison of patients with or without previous back surgery. *Spine.* 1995;20:964-9.

109. Herno A, Airaksinen O, Saari T et al. Lumbar spinal stenosis: a matched-pair study of operated and non-operated patients. *Br.J.Neurosurg.* 1996;10:461-5.

110. Airaksinen O, Herno A, Turunen V et al. Surgical outcome of 438 patients treated surgically for lumbar spinal stenosis. *Spine.* 1997;22:2278-82.

111. Walsh TL, Hanscom B, Lurie JD et al. Is a condition-specific instrument for patients with low back pain/leg symptoms really necessary? The responsiveness of the Oswestry Disability Index, MODEMS, and the SF-36. *Spine.* 2003;28:607-15.

112. Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain.* 1996;65:71-6.

113. Gronblad M, Hupli M, Wennerstrand P et al. Intercorrelation and test-retest reliability of the Pain Disability Index (PDI) and the Oswestry Disability Questionnaire (ODQ) and their correlation with pain intensity in low back pain patients. *Clin.J.Pain.* 1993;9:189-95.

CHAPTER 3
VALIDITY

3.1 Overview

The present thesis research examined validity evidence regarding the use of various measures of walking capacity in LSS populations. Therefore, the concept of validity and its application to the present research have been reviewed. Historically, the three 'types' of validity were content, criterion and construct validity. These distinct types formed the tripartite model of validity.[1;2] However, over time the tripartite model of validity has evolved into a unified model. The current view, supported by a number of the influential writers in this field, is that all validity is construct validity.[3;4] Both Guion [3] and Messick [4] were quoted as saying: "gone is the idea of three types of validity; content and criterion are no more than strands within a cable of a construct validity argument".

3.2 Building a Construct Validity Argument

Construct validity has often been likened to traditional hypothesis testing, which is essentially theory based investigation. Construct validity studies should be structured like an argument, with a series of hypotheses to be supported or disconfirmed.[2;5;6] It has been suggested that each validity argument begin with an explicit definition of the construct of interest.[7] The trait or quality underlying the test is known as the construct of interest and becomes the focus of validity evidence.[1] Once the construct has been defined, one can begin to form a framework of empirical or theory based expected relationships around this construct.[1] One can investigate relationships between measures of the construct of interest and measures of both similar and different constructs. These expected relationships can be used to form the basis of the hypotheses to be tested. Without these articulated theories, there is no construct validity.[6] Construct validity evidence is provided when statistical outcomes provide the support for these hypotheses, that should logically and theoretically be confirmed if a measure is valid.[4] The procedures to be used in construct validity studies are only limited by the creativity of the investigators in formulating hypotheses related to the

construct of interest; any and all lines of validity evidence can be collected to support a validity argument.[6]

To summarize, according to Anastasi [5] what has come to be known as construct validity is actually a comprehensive approach to validity that includes all other recognized validation procedures, including the collection of content and criterion related validity evidence, correlational evidence, as well as evidence of instrument reproducibility. She suggests that almost all information gathered in the process of developing and using a test is relevant to its validity.[5] I believe that the following excerpts from Messick's 1989 influential chapter on validity [4] summarize best the concepts inherent in the unified model of validity:

> *Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on tests scores and other modes of assessment...To validate an interpretive inference is to ascertain the degree to which multiple lines of evidence are consonant with the inference,...Inferences are hypotheses and the validation of inferences is hypothesis testing.[4] (p.13)*

3.3 Types of Validity Evidence

As suggested by Anastasi,[5] under the unified model of construct validity, we can collect other 'types' of validity evidence to support a validity argument, as well as to provide more information about the construct of interest. To support the construct validity argument in the present research both content and criterion related validity evidence was examined.

3.3.1 *Content Validity Evidence*

In examining content validity evidence, a comparison is made between the test or measurement procedure and the construct of interest.[2] According to Cronbach and Meehl,[1] content related evidence demonstrates the degree to which the content of the test samples the subject matter about which conclusions are to be drawn. In examining content validity evidence, we ask if the test or measurement procedure is appropriate for, relevant to and representative of the construct of interest. Generally this type of evidence is collected during the

construction of the test and focuses on the creation of appropriate, relevant and representative items or measures. Judgmental analysis is a replicable procedure used to examine content validity evidence. This analysis involves the selection of multiple judges who are deemed to be specialists in the construct area and who have the experience to determine the relevance and representativeness of a given test in relation to the construct of interest. During this type of analysis the judges are asked to rate the degree of fit between the items or tasks and the construct of interest, as well as to assess the representativeness of the test in relation to the construct of interest. However, in many cases, such judgmental analysis is not conducted, and content validity evidence is based solely on an opinion regarding whether or not the test or procedure 'appears' to be assessing what it is intended to assess.[8]

3.3.2 *Criterion Validity Evidence*

According to Anastasi,[5] the concept of criterion validity evidence dates back to the early 20th century when scientists were suggesting that checking results against a practical external criterion was a means to ensure accuracy. It is assumed that the chosen external variable provides a direct measure of the characteristic or variable of interest, otherwise know as a 'gold standard'.[1;6] The relationship or degree of agreement between the chosen measure and the gold standard is considered criterion related validity evidence. Criterion validity evidence has been broken down into two distinct types: predictive and concurrent. These types of validity evidence are intended to forecast future performance (predictive) or present standing (concurrent) on a criterion variable which was different from those measured in the test.[5] There are a number of different methodologies employed to examine criterion validity evidence. For example, correlational analysis allows us to test hypotheses regarding expected relationships among measures of both similar and different constructs. Convergent validity evidence is provided when measures of constructs that theoretically *should* be related to each other are, in fact, observed to be related to each other, while divergent evidence is provided when measures of constructs that theoretically *should not* be related to each other are observed to not be related to

each other. The multi-method multi-trait matrix described by Campbell and Fiske involves this type of correlational analysis.[9]

3.4 Validity as it Applies to the Current Thesis Research

In a validity investigation, the explicit definition of the construct of interest is of paramount importance, as are the statement and testing of specific hypotheses. Defining the construct of interest for this thesis research, walking capacity in LSS, allowed me to form hypotheses regarding how I believed, based on theory, the outcomes of the various tests of walking in LSS should behave. The aims of this research were to provide information regarding the measures of walking capacity that have been used with LSS populations, and to determine if these measures are indeed providing valid manifestations of the construct of walking capacity in LSS.

Within this thesis research, all 'types' of validity evidence were collected. However, in the spirit of the unified model, they were all essentially serving to provide a greater understanding of the construct of walking capacity and of the meanings and interpretations to be derived from these various tests. Content validity evidence was invoked when suggesting that in the absence of an established criterion, the Self-Paced Walking Test (SPWT) was the best measure of the construct, given that it was a direct operationalization of the construct of walking capacity as defined. Although no official judgmental analysis was conducted, the relevance and representativness of the SPWT for the measurement of walking in LSS was supported by a group of experts, including a spine specialist surgeon, an expert spine researcher, a methodologist and an exercise physiologist. Criterion related evidence was provided when comparing the results of the self-report and treadmill measures to the criterion SPWT, based on hypothesized relationships. Construct validity evidence for both the criterion and surrogate measures was provided in this process, based on hypotheses regarding the expected relationships among the measures of walking capacity and measures of other divergent constructs. The hypotheses regarding the relationships among the measures tested in this research were based on theory related to the construct of interest. Further, these hypotheses were tested using a number of different

statistical methods, including reproducibility and correlational analyses. The evidence acquired through this thesis research is not necessarily convincing for or against the specific use of the tests,[10] but provides a greater understanding of the construct of walking capacity and interpretation of the tests thought to reflect it. This research will hopefully provide impetus for future research into the validity of measures of walking in LSS, given that the both the process of validation and the construct of walking capacity itself should continue to be constantly developed, clarified and refined.[5]

3.5 References

1. Cronbach LJ, Meehl PE. Construct Validity in Psychological Tests. *Psychol.Bull.* 1955;52:281-302.

2. Shepard LA. Evaluating Test Validity. American Educational Research Association. 405-450. 21993. Washingon, DC, Darling-Hammond.
   Ref Type: Report

3. Guion RM. On Trinitarian Doctrines of Validity. *Professional Psychology* 1980;11:385-98.

4. Messick S. Validity. In: Linn RL, ed. *Educational Measurement*. 3 ed. New York: American Council on Education and Macmillan, 1989:13-103.

5. Anastasi A. Evolving Concepts of Test Validation. *Annual Review of Psychology* 1986;37:1-15.

6. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to their Development and Use*. 3 ed. Oxford: Oxford University Press, 2003.

7. Cronbach LJ. Construct Validation after 30 Years. In: Linn RL, ed. *Intelligence: Measurement, Theory and Public Policy*. Champaign IL: University of Illinois Press, 1989:147-71.

8. Mahar MT, Rowe DA. Construct Validity in Physical Activity Research. In: Welk GJ, ed. *Physical Activity Assessments for Health-Related Research*. Champaign IL: Human Kinetics, 2002:51-72.

9. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin* 1959;56:81-105.

10. Cronbach LJ. Five Perspectives on the Validity Arguement. In: Wainer H, Braun HI, eds. *Test Validity*. Hillsdale NJ: Lawrence Erlbaum Associates, 1988:3-17.

CHAPTER 4

STUDY 1

Construct Validity of the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire for the Measurement of Walking Capacity

A version of this chapter has been published in Spine. Tomkins CC, Crites Battié M, Hu R 2007. Construct Validity of the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire for the Measurement of Walking Capacity. Spine. 32(17): 1896-1901.

4.1 Introduction

Degenerative lumbar spinal stenosis (LSS) is a common cause of chronic low back and leg pain in older adults.[1] The primary symptom manifestation of central degenerative LSS is neurogenic claudication, causing pain, numbness, weakness, and tingling in the low back, buttocks, and legs during standing or walking. Generally, such symptoms worsen in neutral or extended positions and improve with lumbar flexion.[2;3] Accordingly, due to pain and discomfort in the lower extremities, people with LSS avoid walking and may have reduced walking capacities.

Measurement of walking capacity in the LSS population is important for the assessment of treatment outcomes, diagnosis, assessment of physical function, and monitoring of the natural progression of the condition. The number of surgeries for LSS has increased drastically in the past few decades and will no doubt continue to increase with the aging population. As such, there will be a growing demand for valid and reliable outcome based measures of walking.[4] Deen et al. suggest that there has been a recent trend toward emphasizing functional status (walking capacity) as a key outcome indicator for patients with LSS symptoms,[5] while Yamashita et al. found self-reported walking difficulty to be the strongest and only independent predictor of patient satisfaction post-surgery for LSS.[4]

A number of self-report instruments have been used to assess walking in persons with LSS, including the low-back specific Oswestry Disability Index (ODI)[6-14], and the Physical Function Scale of the Swiss Spinal Stenosis

Questionnaire.[15-24] Although the ODI is a well-validated instrument for use in the evaluation of lumbar conditions,[1;25-31] it does not address the specific neuro-ischemic characteristics of lumbar spinal stenosis (lower limb numbness, weakness and tingling). The Physical Function Scale, however, was designed specifically for use in the evaluation of physical function in patients with LSS.[1] Although there presently does not exist a 'gold standard' measure of walking capacity in persons with LSS, the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire has been used in a number of studies to specifically evaluate walking in this population.[15-24]

The Swiss Spinal Stenosis Questionnaire was designed to complement existing generic measures of lumbar spine disability and health status in the evaluation of patients with LSS.[1] The three scales used in the questionnaire address symptom severity, satisfaction, and physical function. The five-item Physical Function Scale is used primarily to evaluate walking capacity. These five items assess distance walked, and activities of daily living involving walking. The Physical Function Scale has been used to assess walking as an outcome for surgical and non-surgical treatment in patients with LSS.[15-24]

The psychometric properties of the Physical Function Scale have only been investigated in three studies of which we are aware (Table 4-1). Two of these studies were conducted by Stucki and colleagues, [1;32] the developers of the questionnaire. The three existing studies found the Physical Function Scale to be internally consistent, reliable, and responsive to clinical change.[1;29;32] The only pieces of validity evidence currently available to support the use of the Physical Function scale in the assessment of walking capacity in LSS include the correlations between the Physical Function scale and physicians' assessment of walking capacity ($r=0.47$), as well as the physical dimension of the Sickness Impact Profile (SIP) ($r=0.49$).[1] None of the existing studies were designed to specifically assess the construct validity of the Physical Function scale for the measurement of walking capacity.

Further studies are warranted to examine the construct validity of the Physical Function Scale for use in the measurement of walking capacity in

persons with LSS for a number of reasons. All subjects in the study by Stucki et al.[1] which examined validity of the Physical Function Scale were undergoing surgery for LSS. Validity evidence for the use of the Physical Function Scale has yet to be examined in a group that includes both patients who elect surgery and those that do not, that may represent a broader range of LSS severity. The only construct validity evidence provided to date for the use of the Physical Function Scale in evaluating walking capacity is based on correlations with physicians' subjective assessments of walking capacity, and the physical dimension of the SIP.[1;32] No studies to date have correlated the Physical Function Scale with scores from instruments intended to measure similar constructs, such as the back-pain specific ODI, or other walking-specific questions.

Both the Physical Function Scale [15-24] and the ODI [6-14] have been used to evaluate physical function and walking capacity in patients with LSS. Although the ODI has been validated for use in the measurement of subjective back pain related disability in low-back pain populations,[1;25-29] and specifically for the measurement of back related disability in LSS populations,[30;31] it has not been validated for use in addressing the specific issue of walking capacity in LSS patients. It is not known whether or not the Physical Function Scale and the ODI are highly correlated, and thus comparable in terms of assessing walking capacity in patients with LSS, both for clinical interpretation and comparison of studies related to LSS. No studies to date have reported validity data comparing generic low back scales such as the ODI with instruments designed specifically for the assessment of LSS, such as the Physical Function Scale.[29]

The purpose of the present study was to provide convergent and divergent construct validity evidence for the use of the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire in the measurement of walking capacity, in a heterogeneous LSS population of varying severities, seeking care. This was done by correlating the Physical Function Scale with measures intended to tap similar and different constructs. Construct validity is the extent to which a construct behaves in accordance with the hypotheses concerning how it should behave.[33] Hypotheses regarding the expected magnitude and direction of relationships

between constructs are based on theories, and the testing of these hypotheses provides validity evidence. Because construct validity studies concern associations between abstract concepts, the magnitude of a relationship is judged to be low if correlation coefficients vary between 0.00 and 0.29, moderate between 0.30 and 0.59, and a strong relationship is judged if coefficients are above 0.60.[33-35] Thus, correlation coefficients >0.60 provide convergent validity evidence. The interpretation of correlation coefficients in construct validity studies is less stringent than that used in studies of concurrent or criterion validity, such as when an instrument is compared to an accepted gold standard and much stronger associations are expected (>0.80).[33]

4.2 Materials and Methods

4.2.1 *Subjects*

All subjects in the only previous study investigating the validity of the Physical Function Scale [1] were undergoing surgery for LSS. Conversely, we chose to utilize a more heterogeneous sample of patients seeking care for LSS, to include patients varying degrees of severity. In this way we could validate the Physical Function Scale for use in evaluating walking capacity in patients with LSS who are not necessarily limited to the degree that they elect surgery.

All subjects were >50 years of age, and part of a multi-centre prospective longitudinal study of prognostic factors and outcomes of LSS being conducted in Alberta, Canada. Subjects with suspected lumbar spinal stenosis had been referred by multiple physicians (primarily spine specialist surgeons), to any of four adult imaging facilities during the months of April through October 2004 for lumbar spine imaging. All subjects subsequently had central or combination LSS confirmed on imaging by any of a number of radiologists. Subjects may or may not have been having surgery for LSS. Ethics approval for the study was obtained through the University of Alberta and University of Calgary health research ethics boards.

4.2.2 *Data Collection*

Following review of imaging reports, baseline data were collected using a standardized telephone interview, by a group of trained health sciences students.

Information on participant characteristics of age and gender was acquired as well as history of the current condition, including presence of back and/or leg pain, and duration of back and/or pain. Participants were also asked "are you limited in your ability to walk?" Response options were yes and no. If subjects responded yes they were asked "is this walking limitation due to your back problem?"

The battery of standardized measures included: the Swiss Spinal Stenosis Questionnaire, including the Physical Function subscale to measure walking capacity; [1] Oswestry Disability Index to measure back-related disability; [27] Health Utilities Index Mark 2 and 3 (HUI 2/3), a preference-based measure addressing the issue of health related quality of life; [36] the Centres for Epidemiologic Studies Depression Scale (CES-D) to measure depression; [37] and the Medical Outcomes Social Support Survey (MOS) to examine social support.[38] An additional item to assess walking speed was included from the Oxford Claudication Score.[39] The item addressing pain limited walking distance from the ODI (#4) was isolated for analysis, as well as the HUI3 Single Attribute Utility Score for Ambulation.

4.2.3 *Hypotheses*

The hypothesized relationships supporting construct validity were outlined independent of data collection and prior to analyses. It was hypothesized that the Physical Function Scale would correlate in descending order beginning with the items intended to measure walking capacity specifically (including the item from the ODI addressing pain limited walking distance (#4), the HUI3 Ambulation Utility Score and the walking speed item from the Oxford Claudication Score), followed by the low-back specific ODI, the HUI3 Global Utility Score, the Centres for Epidemiologic Studies Depression Scale (CES-D) and the Medical Outcomes Social Support Survey (MOS). Associations of 0.30-0.59 were judged to provide moderate convergent evidence, >0.60 strong convergent evidence, and >0.80 were very strong convergent evidence. Associations <0.30 were judged to provide divergent validity evidence.[63] However, it was the relative magnitude of the correlations to one another that was most important to the validity argument.

4.2.4 *Data Analysis*

The Physical Function Scale score was calculated as the un-weighted mean of the five items in the scale. The resulting possible scores of 1-4 represent a range from mild to severe limitation in physical function/walking. [1] The item addressing walking distance included the following options: ability to walk "over 3.2km (2mi)", "over 2 blocks but less than 3.2km", "over 15.24m (50ft) but less than 2 blocks", or "less than 15.24m". The other four items in the scale address ability to walk for pleasure, ability to walk for shopping, ability to walk around the house, and ability to walk from bedroom to bathroom. These items are scored using a Likert classification with 4 categories (1= "yes comfortably"; 2= "yes, but sometimes with pain"; 3= "yes, but always with pain"; 4= "no"). The ODI was calculated as a percentage of the total possible score of 53, with a greater score representing greater back related disability.[15;17] The Centres for Epidemiologic Studies Depression Scale and Medical Outcomes Social Support Survey are scored by totaling all item scores, with a higher total indicating greater depression in the CES-D,[37] and greater support in the MOS.[38] The HUI3 Ambulation Score is derived using specific HUI algorithms, with a score ranging from 0 (most disabled) to 1 (no disability).[36] The HUI3 Global score is calculated using 8 single attribute vectors (vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain), on a scale from 0 (dead) to 1(healthy).[36]

Internal consistency was determined for each of the measures using Cronbach's α coefficient. Spearman rank correlation coefficients were used to determine the association between the Physical Function Scale and the other instruments, given that many of the scales are ordinal in nature. An α level of 0.05 was chosen to judge significance. For all correlation coefficients, scatter plots were inspected for linearity and fit statistics were employed. SPSS for Windows, version 15.0 was used for all statistical analysis (SPSS Inc., Chicago, Illinois).

4.3 Results

Of the 75 potential subjects, 3 had data missing in more than one area, and were excluded from the analyses. Thus, as shown in Table 4-2, the study sample consisted of 72 subjects, 51.3% of whom were women. The mean age of the total

sample was 69.5 years. Most subjects (92%) reported a long history of back problems (mean 9.2 years) and the majority (63%) answered yes to both of the following questions regarding walking limitation: "Are you limited in how many minutes you are able to walk continuously?" and "Is this due to your back problem?" The mean scores for all measures can be found in Table 4-2.

Internal consistency of the Physical Function Scale was high (Cronbach's α = 0.88, 95% confidence interval, 0.83-0.92). Inter-item correlations for the Physical Function Scale ranged from 0.54 to 0.87 ($p<0.01$).

The Physical Function Scale correlated >0.60 with the following measures, providing convergent validity evidence: ODI walking specific item (r=0.80), HUI3 Ambulation Score (r=0.62), Oxford Claudication Score walking speed item (r=0.61), ODI total (r=0.72), and the HUI3 Global Utility Score (r=0.61). The Physical Function Scale correlated <0.30 with the Centre for Epidemiologic Studies Depression Scale (r=0.27) and the Medical Outcomes Social Support Survey (r=0.06), providing divergent validity evidence (Table 4-3).

In an additional analysis, we compared subgroups of subjects who had reported walking limitations due to LSS (n=45), and those that did not (n=27). There were no statistically significant differences in the magnitude of correlation coefficients. The hypothesized pattern of correlation coefficients was maintained in both groups.

4.4 Discussion

The mean Physical Function Scale score from the present study of 2.1 suggests slightly less severity on average than that found in previous studies of subjects who were all undergoing surgery for LSS. Thus conclusions from this study are likely to apply most accurately to those patients with Physical Function scale scores in the range of 2.1±0.7. In a study of LSS surgical outcomes, Katz et al. reported pre-operative Physical Function Scale scores ranging from 3.5-3.7 and post-operative scores ranging from 0.9-1.5.[16] In other studies investigating outcome of surgery for LSS, Lee et al. reported a mean pre-operative score of 2.7 and a post-operative score of 2.2,[21] while Sinikallio et al.[40] reported a mean pre-

operative score of 2.5. The mean ODI score in the present study is within the range of scores previously reported for studies of LSS populations.[6;7;14]

The internal consistency of the Physical Function Scale in the present study (α=0.88) is similar to values reported in other studies of LSS conducted by Stucki et al. [1](α=0.82), Katz et al.[15] (α=0.84), and Pratt et al.[29] (α=0.87-0.89).

The hypotheses in the present study regarding the expected relationships between the Physical Function Scale and measures of similar constructs were confirmed, providing convergent construct validity evidence. The Physical Function Scale was correlated highly (r=0.60-0.80) with those items intended to measure walking capacity specifically, including an item from the ODI (#4), which addressed pain limited walking distance, the HUI3 Ambulation Score, and an item from the Oxford Claudication Score regarding walking speed. Based on correlations with the Physical Function Scale, it would appear that the walking item from the ODI is most representative of the construct of walking capacity in LSS (r=0.80). However, generally it is not advisable to use a single item to measure a given construct.

Furthermore, the correlations observed between the other instruments expected to correlate with the Physical Function Scale (ODI, HUI, r=0.60-0.72) provide convergent construct validity evidence for the use of the Physical Function Scale in the assessment of walking capacity in this population. Also as hypothesized, the correlation coefficients associated with the depression measure (CES-D) (r=0.27) and the social support measure (MOS) (r=0.06) were noticeably lower than those of the measures intended to tap constructs similar to the Physical Function Scale, providing divergent validity evidence.

The results of the present study substantiate those of Stucki et al.[1] who provided construct validity evidence for the use of the Physical Function Scale for the assessment of walking in a surgical LSS population, by correlating it with the physical dimension of the SIP (r=0.49), and physicians' assessment of walking capacity (r=0.47). Also, as hypothesized, Stucki et al. found the Physical Function Scale correlated to a greater degree with the physical dimension of the SIP

(r=0.49) than either the global (r=0.43) or psychological (r=0.27) dimensions of the SIP, supporting construct validity.[1]

Both the Physical Function Scale and the ODI have been used to assess walking capacity in persons with LSS.[10;15-24] Given the high correlation between the Physical Function Scale and the ODI in this study (r=0.72), it appears that these instruments are measuring similar constructs. Aside from the correlation between the walking item from the ODI and the Physical Function Scale (r=0.80), the correlation between the ODI total score and the Physical Function Scale was the highest. Unfortunately, due to the nature of this study, we are unable to ascertain if the construct being measured is in fact walking capacity.

There are conflicting opinions regarding the use of the ODI in LSS populations. It has been suggested by Pratt et al.[29] that in assessing LSS, a widely used generic scale for measuring pain related back function, such as the ODI may provide useful information on patient outcomes, as well as allow comparison of treatment interventions and overall function between stenosis patients and those with different low-back conditions. Conversely, Gunzberg et al.[31] suggest that in terms of disability assessment, instruments such as the ODI are not well designed to assess specific disabilities associated with LSS, such as walking problems.

It is apparent that further research is warranted to determine which, if any self-report instrument most accurately represents the actual walking capacity in persons with LSS. If, for example, the ODI were found to be similarly highly correlated with observational measures of walking capacity in patients with LSS compared to the Physical Function Scale, then a disease specific measure such as the Physical Function Scale may not be necessary for the evaluation of this construct.

As previously stated, the aim of the present study was to include LSS patients of varying degrees of severity. In this way, we could validate the Physical Function Scale for use in evaluating walking in those who are not necessarily limited to the degree that they elect surgery. For example, evaluating walking capacity in patients with lesser severity of limitations may be of interest in post-surgical cases, outcomes of conservative treatments, or in longitudinal studies of

the natural history of LSS. However, the inclusion of subjects with no reported walking limitations has the potential to affect the study results. Yet when we conducted an additional analysis comparing subgroups of subjects who had reported walking limitations due to LSS (n=45), and those that did not (n=27), there were no statistically significant differences in the magnitude of correlation coefficients, and the hypothesized pattern of correlation coefficients was maintained, suggesting that results are supportive of construct validity for a full range of self-reported walking limitations. Surprisingly there was variability noted in reported walking ability (distance) even in those who reported no walking limitation, with scores for the two items intended to address walking distance (Physical Function Scale #1 and ODI #4) ranging from 1-4 and 1-6 respectively. The fact that both subgroups reported a range of walking distances implies that the perception of a presence or absence of a walking limitation may not represent actual walking capacity. Further research is warranted to look specifically at how observational measures of walking capacity relate to self-reported limitations in persons with LSS.

4.5 Conclusions

In summary, both convergent and divergent validity evidence were provided to support the use of the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire in the assessment of walking capacity in persons with LSS. Although high correlations were found between the Physical Function Scale and those items and instruments intended to measure walking capacity in LSS, including the back-pain specific ODI, it cannot be ascertained from the present study that the construct being measured is in fact walking capacity. Further research is warranted to investigate criterion validity evidence for the use of the Physical Function Scale, and the ODI in the measurement of walking capacity in LSS, by examining the relationships between these self-report instruments and observational measures of walking.

Table 4-1. Psychometric Properties of the Physical Function (PF) Scale of the Swiss Spinal Stenosis Questionnaire

	Source	Study details	Test statistic	Result
Internal Consistency	Stucki et al. [1]	N=193	Cronbach's Coefficient α	PF scale =0.82
	Pratt et al. [29]	N=29	Cronbach's Coefficient α	PF scale=0.87-0.89
Convergent Validity	Stucki et al. [1]	N=193 Physical function scale correlated with other measures of physical function	Spearman's Convergent Correlation Coefficient	Correlation of the PF scale with: Physical SIP (r = 0.49)* Global SIP (r= 0.43)* Physician's assessment of walking capacity (r=0.47)*
Test Re-test Reliability	Stucki et al. [1]	N=23 Test-retest method, 2 weeks apart	Spearman's Correlation Coefficient	PF scale=0.94* Walk for shopping=0.71* Walking distance=0.93*
	Pratt et al. [29]	N=29	Intraclass Correlation Coefficient (ICC)	PF scale=0.82
Responsiveness	Stucki et al. [1]	N=130	Standardized Response Mean	PF scale = 1.07 Walking distance= 0.56 Walking outdoors= 1.01
	Stucki et al. [1]	N=130 Correlation between change score for the PF scale and satisfaction score on the Swiss Spinal Stenosis Measure	Spearman's Correlation Coefficient	R=0.72* (95 % confidence interval, 0.63-0.79)
	Stucki et al. [32]	N=130 Compared responsiveness of PF scale to the SIP and the Roland Morris Scale	Standardized Response Mean	PF Scale=1.07 Roland= 0.77 SIP = 0.69 Physical SIP= 0.62

All subjects for both studies were taken from an ongoing prospective multi-centre observational study of patients undergoing surgery for LSS. Mean age of subjects for both studies was 68 [1; 32].
*p<0.01; SIP: Sickness Impact Profile.

Table 4-2. Characteristics of the Study Group (n=72)

Variable	N	Percentage or Mean±SD
Age (yrs)	72	69.5±11.3
Gender		
➢ Female	37	51.3
Back pain present	66	92
Duration of back pain (yrs)		9.2±10.9
Leg symptoms present	61	84.7
Duration of leg pain (yrs)		7.5±8.9
Both back and leg symptoms present	56	77.8
Walking limited due to back problem	45	62.5
Outcome Measures		
Physical Function Scale (1-4)	72	2.1±0.7
ODI (%)	72	43.1±13.7
HUI3 Global (0-1)	72	0.63±0.3
HUI3 Ambulation Score (0-1)	72	0.83±0.2
MOS (0-100)	72	73.7±15.6
CES-D (0-100)	72	17.9±6.8

Table 4-3. Convergent and Divergent Correlation Coefficients for the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire in the Evaluation of Walking in Subjects with Lumbar Spinal Stenosis (n=72)

Spearman's Rho (n=72)	PF SCALE	ODI	ODI WALK ITEM	OCS ITEM	HUI3 GLOBAL	HUI3 AMB	CES-D	MOS
PF SCALE	**1.00**							
ODI	0.72**	**1.00**						
ODI WALK	0.80**	0.72**	**1.00**					
OCS SPEED	0.60**	0.55**	0.66**	**1.00**				
HUI3 GLOBAL	0.61**	0.69**	0.53**	0.57**	**1.00**			
HUI3 AMB	0.62**	0.64**	0.68*	0.62**	0.68**	**1.00**		
CES-D	0.27*	0.53**	0.31*	0.34**	0.45**	0.42**	**1.00**	
MOS	0.06	0.09	0.04	0.11	0.16	0.05	0.21	**1.00**

** $p<0.01$ *$p<0.05$ Gray shading indicates divergent correlation coefficients, while no shading indicates convergent coefficients. PF Scale = Physical Function Scale of the Swiss Spinal Stenosis Questionnaire; ODI = Oswestry Disability Index; ODI WALK = Walking item from the ODI addressing pain limited walking distance; OCS SPEED = Walking item from the Oxford Claudication Score; HUI3 GLOBAL = Health Utilities Index Mark 3 Global Health Related Quality of Life Utility Score; HUI3 AMB = Single Attribute Utility Score for Ambulation; CES-D = Centre for Epidemiologic Studies Depression Scale; MOS = Medical Outcomes Social Support Survey.

4.6 References

1. Stucki G, Daltroy L, Liang MH et al. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine.* 1996;21:796-803.

2. Katz JN, Dalgas M, Stucki G et al. Degenerative lumbar spinal stenosis. Diagnostic value of the history and physical examination. *Arthritis Rheum.* 1995;38:1236-41.

3. Katz JN, Dalgas M, Stucki G et al. Diagnosis of lumbar spinal stenosis. *Rheum.Dis.Clin.North Am.* 1994;20:471-83.

4. Yamashita K, Hayashi J, Ohzono K et al. Correlation of patient satisfaction with symptom severity and walking ability after surgical treatment for degenerative lumbar spinal stenosis. *Spine.* 2003;28:2477-81.

5. Deen HG, Jr., Zimmerman RS, Lyons MK et al. Test-retest reproducibility of the exercise treadmill examination in lumbar spinal stenosis. *Mayo Clin.Proc.* 2000;75:1002-7.

6. Herno A, Airaksinen O, Saari T et al. Surgical results of lumbar spinal stenosis. A comparison of patients with or without previous back surgery. *Spine.* 1995;20:964-9.

7. Herno A, Airaksinen O, Saari T et al. Lumbar spinal stenosis: a matched-pair study of operated and non-operated patients. *Br.J.Neurosurg.* 1996;10:461-5.

8. McGregor AH, Hughes SP. The evaluation of the surgical management of nerve root compression in patients with low back pain: Part 1: the assessment of outcome. *Spine.* 2002;27:1465-70.

9. Herno A, Saari T, Suomalainen O et al. The degree of decompressive relief and its relation to clinical outcome in patients undergoing surgery for lumbar spinal stenosis. *Spine.* 1999;24:1010-4.

10. Shabat S, Leitner Y, Nyska M et al. Surgical treatment of lumbar spinal stenosis in patients aged 65 years and older. *Arch.Gerontol.Geriatr.* 2002;35:143-52.

11. Herno A, Airaksinen O, Saari T. Computed tomography after laminectomy for lumbar spinal stenosis. Patients' pain patterns, walking capacity, and subjective disability had no correlation with computed tomography findings. *Spine.* 1994;19:1975-8.

12. Herno A, Partanen K, Talaslahti T et al. Long-term clinical and magnetic resonance imaging follow-up assessment of patients with lumbar spinal stenosis after laminectomy. *Spine.* 1999;24:1533-7.

13. Tenhula J, Lenke LG, Bridwell KH et al. Prospective functional evaluation of the surgical treatment of neurogenic claudication in patients with lumbar spinal stenosis. *J.Spinal Disord.* 2000;13:276-82.

14. Airaksinen O, Herno A, Turunen V et al. Surgical outcome of 438 patients treated surgically for lumbar spinal stenosis. *Spine.* 1997;22:2278-82.

15. Katz JN, Lipson SJ, Chang LC et al. Seven- to 10-year outcome of decompressive surgery for degenerative lumbar spinal stenosis. *Spine.* 1996;21:92-8.

16. Katz JN, Lipson SJ, Lew RA et al. Lumbar laminectomy alone or with instrumented or noninstrumented arthrodesis in degenerative lumbar spinal stenosis. Patient selection, costs, and surgical outcomes. *Spine.* 1997;22:1123-31.

17. Iversen MD, Katz JN. Examination findings and self-reported walking capacity in patients with lumbar spinal stenosis. *Phys.Ther.* 2001;81:1296-306.

18. Katz JN, Stucki G, Lipson SJ et al. Predictors of surgical outcome in degenerative lumbar spinal stenosis. *Spine.* 1999;24:2229-33.

19. Frazier DD, Lipson SJ, Fossel AH et al. Associations between spinal deformity and outcomes after decompression for spinal stenosis. *Spine.* 1997;22:2025-9.

20. Simotas AC. Nonoperative treatment for lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;153-61.

21. Lee J, Hida K, Seki T et al. An interspinous process distractor (X STOP) for lumbar spinal stenosis in elderly patients: preliminary experiences in 10 consecutive cases. *J.Spinal Disord.Tech.* 2004;17:72-7.

22. Hansraj KK, O'Leary PF, Cammisa FP, Jr. et al. Decompression, fusion, and instrumentation surgery for complex lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;18-25.

23. Hansraj KK, Cammisa FP, Jr., O'Leary PF et al. Decompressive surgery for typical lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;10-7.

24. Simotas AC, Dorey FJ, Hansraj KK et al. Nonoperative treatment for lumbar spinal stenosis. Clinical and outcome results and a 3-year survivorship analysis. *Spine.* 2000;25:197-203.

25. Walsh TL, Hanscom B, Lurie JD et al. Is a condition-specific instrument for patients with low back pain/leg symptoms really necessary? The responsiveness of the Oswestry Disability Index, MODEMS, and the SF-36. *Spine.* 2003;28:607-15.

26. Gronblad M, Hupli M, Wennerstrand P et al. Intercorrelation and test-retest reliability of the Pain Disability Index (PDI) and the Oswestry Disability Questionnaire (ODQ) and their correlation with pain intensity in low back pain patients. *Clin.J.Pain.* 1993;9:189-95.

27. Fairbank JC, Couper J, Davies JB et al. The Oswestry low back pain disability questionnaire. *Physiotherapy.* 1980;66:271-3.

28. Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain.* 1996;65:71-6.

29. Pratt RK, Fairbank JC, Virr A. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. *Spine.* 2002;27:84-91.

30. Herno A, Airaksinen O, Saari T et al. Computed tomography findings 4 years after surgical management of lumbar spinal stenosis. No correlation with clinical outcome. *Spine.* 1999;24:2234-9.

31. Gunzburg R, Keller TS, Szpalski M et al. Clinical and psychofunctional measures of conservative decompression surgery for lumbar spinal stenosis: a prospective cohort study. *Eur.Spine J.* 2003;12:197-204.

32. Stucki G, Liang MH, Fossel AH et al. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J.Clin.Epidemiol.* 1995;48:1369-78.

33. Gross DP, Battie MC. Construct validity of a kinesiophysical functional capacity evaluation administered within a worker's compensation environment. *J.Occup.Rehabil.* 2003;13:287-95.

34. Innes E, Straker L. Validity of work-related assessments. *Work.* 1999;13:125-52.

35. Domholdt E. *Physical Therapy Research: Principles and Applications*. 2nd ed. ed. Philadelphia, PA: Saunders, 2000.

36. Feeny D, Furlong W, Boyle M et al. Multi-attribute health status classification systems. Health Utilities Index. *Pharmacoeconomics.* 1995;7:490-502.

37. Radloff LS. The CES-D scale: A self-report depression scale for research in general populations. *Appl.Psychol.Meas* 1977;1:385-401.

38. Stewart AL, Ware JE. *Measuring Function and Well Being: The Medical Outcomes Approach*. Rand Corporation: 1991.

39. Makan P, Fairbank J, Wanders L. Clinical Assessment of Lumbar Spinal Stenosis. *Journal of Bone and Joint Surgery (Br)* 1998;30:158.

40. Sinikallio S, Aalto T, Airaksinen O et al. Somatic comorbidity and younger age are associated with life dissatisfaction among patients with lumbar spinal stenosis before surgical treatment. *Eur.Spine J.* 2006.

CHAPTER 5

STUDY 2

Reproducibility of Measures of Walking Capacity in Lumbar Spinal Stenosis

5.1 Introduction

Valid and reproducible measurement of walking capacity is important for the assessment of both treatment outcomes and the natural progression of the condition in patients with lumbar spinal stenosis (LSS). There has been a recent trend toward emphasizing walking capacity as a key outcome indicator for patients receiving both surgical and non-surgical treatment for LSS.[1] Accordingly, most of the studies published recently that focused on treatment for LSS reported some measurement of walking capacity.[2-7] In addition to use as an outcome indicator for treatment, it has been suggested that measurement of walking using treadmill tests may be useful in the differential diagnosis of LSS.[8] Thus, there is a growing demand for valid and reproducible measures of walking capacity in LSS populations.[9]

Walking capacity has been defined in the present research as the distance a person with LSS is able to walk without support on a level surface at a self-selected speed before being forced to stop due to symptoms of LSS. There are a number of self-report and observational measures currently available for use with LSS patients that address to a certain degree the construct of walking capacity, as defined above. However, to date there is no 'gold standard' or criterion measure of walking capacity for this population. Measures that have been used often to evaluate walking in LSS include the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire (Physical Function Scale),[10] the Oswestry Disability Index (ODI),[11] and various treadmill protocols.[1;8;12-22] There is some validity evidence available to support the use of the Physical Function Scale,[10;23;24] the ODI,[8;14] and treadmill testing [14] in the evaluation of walking capacity in LSS. However, it is not certain that the construct being tapped using any of these measures, is, in fact, walking capacity. Further research is warranted to investigate criterion validity evidence for the use of the Physical Function Scale, the ODI and treadmill tests in the measurement of walking capacity in LSS by examining the relationships

between these instruments and a construct specific observational measure of walking.

In order to examine criterion validity evidence regarding measurement of walking capacity in LSS, a criterion measure was developed based on the aforementioned operational definition. The proposed criterion measure, the Self-Paced Walking Test (SPWT) asks LSS patients to walk at their own pace on a level surface (indoor track) until forced to stop due to symptoms of LSS. This test has some evidence supporting its construct validity as it is a direct operationalization of the construct of interest and is relevant to and representative of the construct of walking capacity as defined. However, before validity evidence can be examined using the SPWT, it was essential that its reproducibility be investigated. Reproducibility concerns the degree to which repeated measurements yield similar results.[25] In order to be able to detect changes in the walking capacity of patients with LSS, we first needed to determine whether the SPWT provided similar results over a relatively short period of time when no change was expected.

Given that the SPWT is a new test, its reproducibility has yet to be examined. However, a number of authors have examined the reproducibility of the Physical Function Scale in LSS populations. Stucki et al.[10] examined test re-test reproducibility in 23 patients undergoing surgery for LSS over an interval of two weeks. They found the Physical Function Scale to be reproducible, with a reported Spearman correlation between test administrations of 0.94. Pratt et al.[26] reported an ICC of 0.82 over an interval of one week in a study with 29 LSS patients, while Thornes et al. [24] reported an ICC of 0.89 over an interval of one week in a study with 75 LSS patients post-surgery. High reproducibility has also been reported for the ODI in numerous studies with low back pain populations.[27;28] However, only one study of which we are aware has examined reproducibility of the ODI in an LSS population. Pratt et al. reported an ICC of 0.89 for the ODI over a test re-test interval of one week.[26]

The primary purpose of the present study was to examine the reproducibility of the Self-Paced Walking Test. A secondary goal was to assess

reproducibility of the surrogate measures of walking capacity in LSS, including the Physical Function Scale, the ODI and a construct specific question regarding walking capacity.

It was hypothesized that walking capacity would be stable over a period of up to three weeks, demonstrated by high agreement between two administrations of the SPWT and the surrogate measures of walking capacity in LSS. Specifically, it was hypothesized that there would be no significant differences between administrations of the tests and that all intra-class correlation coefficients (ICC) would be >0.80. We also examined the mean absolute deviation (MAD) between testing occasions to provide more information regarding agreement between the values of the two measurements for each test.

5.2 Materials and Methods

5.2.1 *Subjects*

All subjects had been referred to one of five spine specialist surgeons in the Edmonton region with suspected LSS. At the time of their consultation with their spine specialist surgeon, subjects who met the inclusion criteria were given information about the study. The patients who expressed an interest in the study were later contacted by telephone and provided with more information. The inclusion criteria were >45 years of age with central or combination LSS confirmed on imaging (MRI/CT) and by a spine specialist surgeon. All subjects also had self-reported LSS associated walking limitations or symptoms exacerbated by walking (neurogenic claudication). Exclusion criteria included surgery for LSS within the past year or any co-morbid conditions that would limit walking capacity or make exercise medically inadvisable as judged by the subjects' physician.

5.2.2 *Data Collection*

Subjects were asked to come to the testing facility three times, with a maximum of 7 days between visits. Over the course of three visits subjects were asked to complete the Self-Paced Walking Test (SPWT) twice and a treadmill protocol once. Subjects were randomized to performing either the treadmill test or the criterion SPWT during the first visit. They completed the other measure at the

time of the second visit. Those who participated in a third visit completed the second SPWT. Given that we only conducted one treadmill test we did not examine reproducibility of the measurements obtained with the treadmill protocol.

During the first two visits, subjects also completed a battery of questionnaires, including the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire and the Oswestry Disability Index. Subjects were also asked the following construct specific question: "If you were to go for a walk today, how far would you be able to walk at your own pace, on level ground before being forced to stop due to symptoms of lumbar spinal stenosis (metres)?" In order to provide a distance reference to all subjects during the completion of the walking distance questions, subjects were informed that the distance on the straight-away of the track that they were able to view was 60m (~200ft).

In order to assess perceived changes in walking capacity between the test re-test administrations of the SPWT, subjects were asked directly before their second SPWT to rate change in their walking capacity on a 7-point Likert scale with the following options: a great deal better, moderately better, a little bit better, about the same, a little bit worse, moderately worse and a great deal worse.

If subjects were taking medications, they were asked to be consistent with their medications on the testing days. In addition, they were asked to do their best to stay consistent with any other treatments they were receiving between visits, to minimize any changes in their condition.

Ethics approval for the study was obtained through the University of Alberta health research ethics board.

5.2.3 *Self-Paced Walking Test Protocol*

Subjects were instructed to walk continuously at their own pace around the outer lane of the track until they felt they had to stop due to symptoms of LSS (or other reasons) or until the time limit of 30 minutes had been reached. At no time were subjects encouraged or prompted to continue. Subjects were also asked to indicate when they first experienced a change in symptoms and to describe the nature of the symptoms. Test termination was defined as a complete stop of 3 seconds. Throughout the test, the administrator followed at a comfortable

distance, with a rolling wheel instrument to measure distance (Lufkin® Pro-Series Model PSMW38) and a stopwatch. The following information was collected: total distance and time walked, distance and time to onset of symptoms, the nature and location of symptoms (pain, numbness/tingling, weakness or fatigue), average walking speed and the reason for test termination should they not walk for the full 30 minutes (symptoms of LSS, fatigue, shortness of breath, dizziness, pain or discomfort due to co-morbidities).

5.2.4 *Data Analysis*

The Physical Function Scale score was calculated as the un-weighted mean of the five items in the scale. The resulting possible scores of 1-4 represent a range from mild to severe limitation in physical function/walking.[10] The item from the Physical Function Scale addressing walking distance was isolated for analysis, and measured on a scale from 1-4 which included the following options: ability to walk "over 3 km" (1), "over 250m but less than 3 km" (2), "over 15m but less than 250m" (3), or "less than 15m" (4).[5] The ODI was calculated as a percentage of the total possible score of 53, with a greater score representing greater back related disability.[11] The item addressing pain limited walking distance from the ODI was also isolated for analysis, and measured on a scale from 1-6 which included the following options: "I am in bed most of the time" (6), "I can only walk using a stick or crutches" (5), "Pain prevents me from walking more than 500m" (4), "Pain prevents me from walking more than 1km" (3), "Pain prevents me from walking more than 2km" (2), " Pain does not prevent me from walking any distance" (1).

Descriptive statistics were used to describe the characteristics of the sample. Paired t-tests were used to assess differences between testing occasions for all measures. An alpha (α) level of 0.05 was chosen to judge significance. Reproducibility of the SPWT and the self-report instruments were statistically evaluated using intra-class correlation coefficients (ICC – two way mixed model). The ICC is intended to be used when analyzing repeated measures data for the same variable. In contrast to Pearson Product Moment Correlations, which are often used to report test re-test reliability, the ICC is sensitive to changes in the

order and magnitude of the means between repeated measures. Agreement between testing occasions was also examined using mean absolute deviation (MAD), which is the sum of the absolute differences between testing occasions divided by the number of subjects. A test with perfect reproducibility would have a MAD of zero. The MAD was employed because it is not affected by the direction (+/-) of the difference.

In order to determine whether a change in walking capacity (distance) on the SPWT could be predicted by patients' perceived changes in their walking capacity, change scores were correlated with actual change in walking distance between the two SPWTs.

In anticipation of a possible ceiling effect due to some patients walking for the full 30 minutes, subjects who walked for 30 minutes during either SPWT were removed post-hoc from the sample and all analyses were repeated.

5.3 Results

Of the 41 subjects who participated in this study, 28 returned for both the second and third visits. Thus the study sample for examining measurement reproducibility consisted of 28 subjects with a mean age of 66.8±7.7 years, 16 of whom were women. Two subjects reported walking routinely with a walker and three with a cane. The mean number of days between SPWTs was 10.7 days ±4.9 days with a range of 4-21 days. The mean walking time for the first SPWT was 15.2 minutes ±11.2 minutes and 14.7 minutes ±11.1 minutes for the second SPWT.

The primary variable of interest for this study was walking distance measured in the SPWT. The other walking capacity measures which were investigated included SPWT distance to first symptoms and walking speed, as well as the mean scores for the Physical Function Scale Score, Physical Function Scale walking distance item, the ODI, ODI walking distance item and self-predicted walking distance. Mean scores for the two administrations of the walking capacity measures, as well as the range of scores are presented in Table 5-2. The mean differences between testing occasions, as well as the mean absolute deviations (MAD) are also presented in Table 5-2.

5.3.1. *SPWT Total Walking Distance*

As hypothesized, there was no significant difference between testing occasions for total walking distance ($p>0.05$). Table 5-1 demonstrates the actual difference in distance observed for each subject over the two tests, ranked from smallest to largest percent difference (difference in distance between tests as a percent of the first SPWT distance). The mean difference between the two SPWT testing occasions was 39.0m, which was not found to be significantly different from zero ($p>0.05$). However, the MAD was 151.5 m. This higher value for the MAD can be explained by the presence of some differences that were large. According to Table 5-1, the absolute differences ranged from zero to 533.7m. Subjects can be broken down into six groupings, based on absolute differences in distance. Ten subjects had a difference between tests of less than 60m (range of 0m-53.6m), six subjects had a difference between 60m and 150m (range of 89.1m-138.1m), two subjects had a difference between 150m and 200m (range of 173.5m-179.8m), six had a difference between 200m and 300m (range of 222.8m-259.4m), two had a difference between 300m and 400m (range of 300.6m-338.6m) and two had differences greater than 400m (440.1m and 533.7m). This progression of differences resulted in a mean absolute deviation of 151.5m. Table 5-1 also presents individual differences between SPWTs expressed as a percentage of the first SPWT. The differences range from 0% to 743.2%. Of the 28 subjects, 14 (50%) had a difference between SPWTs of less than 10%, five subjects (18%) had a difference between 10% and 25% and seven subjects (25%) had a difference between 25% and 50%. Only two subjects (7%) had a difference between SPWTs of more than 50%. These two subjects walked 92.0% and 743.2% further during the second SPWT.

It is also apparent from Table 5-1 that there is no systematic relationship between the mean distance walked over the two tests and the difference in distance between the tests. Accordingly, the correlations between the mean absolute deviations (MAD) of the subjects and the distances they walked at time 1 and time 2 were low ($r=0.18$ and $r=0.08$, respectively). Of the 28 subjects, 13 (46%) walked further during the first SPWT, 14 (50%) walked further in the

second SPWT, and one subject was unchanged, suggesting that a learning phenomenon was likely not associated with the use of this test.

As hypothesized, the ICC for total distance walked in the SPWT (0.98) exceeded the value of 0.80. This indicates that despite the differences in the total walking distances observed between first and second SPWT, the subjects were ranked essentially the same on first and second occasion (Table 5-3).

5.3.2 *SPWT Distance to First Symptoms and Walking Speed*

No significant differences were observed between assessments for SPWT distance to first symptoms or walking speed ($p>0.05$). As hypothesized, the ICCs for walking speed (0.95) and distance to first symptoms (0.80) were >0.80, demonstrating good reproducibility (Table 5-3). The MAD for walking speed was 0.25km/h, while the MAD for distance to first symptoms was 214.9m (Table 5-2).

5.3.3 *Self-Report Measures*

The mean number of days between the two administrations of the self-report measures was 7.1±2.7 with a range of 2 to14 days. There was no significant difference between assessments for any of the walking measures ($p>0.05$) (Table 5-2). As hypothesized, the ICC for the ODI (0.90) was >0.80, as was the ICC for the ODI walking distance item (0.89). The ICC for the Physical Function Scale (0.77) and its walking distance specific item (0.74) were slightly lower than expected, while the ICC for self-predicted walking distance was much lower than expected (0.50) (Table 5-3). The mean absolute deviation (MAD) for the Physical Function Scale and its walking distance specific item were 0.37 and 0.32, respectively. The MAD for the ODI and its walking distance specific item were 4.1% and 0.4% respectively. The MAD for the self-predicted walking distance was 837.6m.

5.3.4 *Change Score*

On the whole, the subjects' perceived changes in walking capacity were not correlated with observed changes between the two SPWTs. This was evidenced by a lack of relationship between the walking capacity change score and the actual change in walking distance over the two SPWTs ($r=0.12$, $p>0.05$). Furthermore, the change score was not correlated with a change in pre-SPWT

pain between tests (measured on a visual analog scale) ($r=0.28$, $p>0.05$), nor a change in the Physical Function Scale ($r=0.28$, $p>0.05$), the ODI ($r=0.23$, $p>0.05$) or self-predicted walking distance ($r=0.23$, $p>0.05$).

5.3.5 *Post-hoc Analyses*

Post-hoc analyses were conducted after removing any subjects who walked for the full 30 minutes, in order to eliminate the possibility of a ceiling effect. Of the 28 patients who completed two SPWTs, eight (28%) walked for the full 30 minutes. Similar results were obtained for all analyses with these subjects removed. All subjects who did not walk for the full 30 minutes ($n=20$) reported terminating the test due to symptoms of LSS. The mean walking distance for these subjects was 586.8m±416.5, with a mean difference in distance between SPWTs of 44.8m and an ICC for total distance walked of 0.87 (0.69-0.95). The MAD for total walking distance in this group was 159.0m. No significant difference was observed between testing occasions for any of the walking measures.

Interestingly, all eight subjects who walked for the full 30 minutes also reported symptoms of neurogenic claudication during the test. The mean total distance for these eight subjects was 2190.5m±616.7, with a mean difference in distance between SPWTs of 24.6m and an ICC for total distance walked of 0.78 (95% confidence interval, 0.23-0.95). The MAD for total walking distance was 132.7m. No significant difference was observed between testing occasions for any of the walking measures.

An additional post-hoc analysis was conducted in order to determine whether the duration of time between tests affected the agreement between tests for the full sample of 28 subjects. No relationship was observed between the number of days between SPWTs and the difference in distance walked between tests ($r=0.13$, $p>0.05$).

Because it is possible that some variation in repeat measurements may have occurred due to day to day variations in a subject's condition, an additional analysis was done using data from only those who reported no change in their perceived walking capacity between SPWTs (Table 5-4). The results of the 15

subjects represent to some degree the natural variation in walking capacity that could be expected over a period of 2-days to 3 weeks, in those subjects who do not detect a change. The mean difference between SPWTs for this group for total walking distance was 24.8 m ± 222.5 (2% of mean total distance). The individual absolute differences between tests ranged from an increase in walking distance of 92.0% to a decrease in distance of 36.1%.

5.4 Discussion

Results of the present study demonstrate reproducibility of the Self-Paced Walking Test, the ODI and the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire for the measurement of walking capacity in a sample of patients with LSS. These results are important given the recent emphasis on functional status and walking capacity in studies of LSS diagnosis and treatment outcomes. Many patients with LSS, including those in the present study, report that their symptoms vary from day to day, which may cause variation in response to testing of walking capacity.[26] These variations in patient symptoms may affect the reproducibility of a test.[26] Thus, studying reproducibility in these measures was necessary to ensure that the error involved in measuring walking capacity in this population is small enough to detect actual changes in patients' walking.[29]

5.4.1 *Self-Paced Walking Test as a Criterion Measure of Walking Capacity in LSS*

The primary purpose of the present study was to provide evidence regarding the reproducibility of the proposed criterion measure of walking capacity in LSS, the SPWT. The SPWT is a direct operationalization of the construct of walking capacity in LSS, defined as the distance a person with LSS is able to walk without support on a level surface at a self-selected speed before being forced to stop due to symptoms of LSS. This test is thought to have construct validity as it is a direct operationalization of the construct of walking capacity as defined. This test is also thought to closely replicate 'normal' or authentic walking conditions in patients with LSS, within a standardized environment. Total distance was selected as the construct of interest as opposed to total time, given that walking is usually goal oriented, with patients needing to reach a destination of a particular distance.[12]

In the present study, all aspects of the SPWT were found to be reproducible, including total walking distance (construct of interest), as well as walking speed and distance to first symptoms. Reproducibility evidence from the present study supports the use of the SPWT as a criterion in examining validity evidence for various other measures of walking capacity in LSS. It is essential that the SPWT be reproducible for use as a criterion measure so that associations with other measures of walking are not attenuated by low reliability of the criterion.

5.4.2 *Total Walking Distance in the SPWT*

Total distance measured in the SPWT was found to be highly reproducible at the group level with an ICC of 0.98 (95% confidence interval, 0.95-0.99) and a MAD of 151.1 m. At the group level, the MAD is relatively small (<15%) considering the mean overall distance of 1142 m, and the wide range of walking distances included in the sample (31.7-3077.0m). The MAD corresponds to 4.9% of the range for the first SPWT and 5.3% for the second SPWT. These results are consistent with studies examining the natural history of LSS, which suggest that walking capacity in LSS remains largely unchanged for up to 10 years.[30] These results are also similar to those reported by Deen et al.[1] and Moon et al.[22] in the only studies which have examined reproducibility of observational tests for the measurement of walking capacity in LSS.[1] Using a treadmill protocol, Deen et al. reported a concordance correlation coefficient of 0.98 between testing occasions for total walking time in a Self-Paced trial,[1] while Moon et al. reported a Spearman correlation of 0.92 for total time between re-tests of the same Deen protocol, at 1.2mph.[22] Test re-test intervals for these two studies ranged from 1-4 days.[1;22]

When examining agreement between testing occasions at the individual level however, total walking distance was found to be less reproducible, with some individuals showing substantial instability in walking capacity. Fifty percent of subjects had a difference between SPWTs of less than 10%, suggesting that the distance measured in the SPWT is relatively stable and reproducible in a large portion of subjects over a period of four days up to three weeks. However, 32%

had changes greater than 25%. This begs the question, what is the source of this observed variation at the individual level? Is the walking capacity of subjects actually changing over the time between tests, or is the observed change due to measurement error? Unfortunately, it is not possible to know for certain. Many possible sources of error were controlled. The protocol was standardized and consistent for all subjects. The tests were conducted by the same person, with the same measurement tool, in the same location for all subjects. The majority of tests were conducted at the same time of day. Subjects were asked to wear the same shoes for both tests, as well as be consistent with any medications or pre-test activities. Thus, the changes observed between tests likely reflect actual changes in intrinsic factors, such as motivation or walking capacities. If indeed the changes between tests are due to actual changes in walking capacity, future research is warranted to investigate what factors might predict such changes over a short period of time, such as symptom severity and motivation.

The results of the present study bring into light the need for parameters of both high reliability (ICC), and agreement (MAD). While the high ICC of 0.98 suggests that the SPWT is highly reproducible, on examining actual agreement between the tests using the MAD, we see that a number of subjects had differences between tests that were substantial. The high ICC suggests that subjects were ranked the same on the two tests, but does not adequately reflect the actual difference between tests for individuals. It has been suggested that ICCs are most appropriate for examining reproducibility of measures used for discriminative purposes.[31] That is, the measurement error should be small compared to the variability between persons among whom the instrument needs to distinguish.[25] Therefore, if the differences *between* subjects are large, more measurement error is acceptable. This explains in part why the ICC is quite high in the present study. The reason being, that the range in walking distances observed and differences between subjects was quite large in comparison with the measurement error. However, when you are interested in evaluating change on an individual level, measures of agreement, such as the MAD are warranted. In the case of agreement, the variability between people does not matter, only the

measurement error is important.[25] Therefore, given that a test such as the SPWT could potentially be used for both discriminative and evaluative purposes, measures of both reliability and agreement are warranted.

When examining reproducibility at the individual level, the error in testing should be smaller than the changes you wish to detect. Unfortunately, based on the current research we cannot determine what constitutes a significant change in walking capacity for LSS patients. What we can provide, is information regarding the possible variation in walking distance (and other walking capacity measures) observed in a group of LSS subjects over a period of four days up to three weeks. Specifically, if we take those subjects who reported no change in walking capacity between tests, we would expect a minimal amount of change. These subjects are the closest we have to a group who consider their own walking to be stable. Therefore, you would expect that the changes observed between tests for these individuals would reflect natural fluctuations that are not considered to be significant by the patients themselves (Table 5-4). For total walking distance in the SPWT, the mean difference between tests for this group was 24.8m ± 222.5 (2% of mean overall walking distance). The percentage difference between tests ranged from -36.1% to 92%. This result suggests that even in those patients who perceived no change in their walking capacity between test days, substantial fluctuation in walking capacity up to 92% may be observed for some patients over a short period of time. The maximum differences between repeat administrations of the other walking capacity measures (calculated as the difference divided by the value obtained during the first test) were 96.0% for distance to first symptoms in the SPWT, 22.6% for walking speed, 50% for the Physical Function Scale, 27% for the ODI and 83.3% for self-predicted distance. Therefore, given the large differences observed between tests for some subjects, if accuracy is of great importance in measurement of walking distance, multiple measurements should be taken.

There were a few subjects in the present study who *were* able to detect a change in their walking capacity between tests. In the question addressing change in walking capacity since the previous SPWT, two subjects reported that their

walking capacities were "a little bit worse" compared to their first SPWT. These subjects were identified as having absolute differences between SPWTs that were noticeably larger than the other subjects. The differences between test administrations for these subjects were 545m and 440m, respectively, indicating that there was likely a change in their walking capacity between testing occasions. Both subjects walked further (27% and 36%, respectively) during the first test, suggesting that their symptoms had potentially worsened over the 14 to 15 days between tests. Examination of visual analog pain scales confirmed that pre-walking pain had increased in both subjects over the time between SPWTs. The results for these two individual cases confirm again that although walking appears to be relatively stable in most patients with LSS, there are patients whose walking capacity may fluctuate over a short period of time due to acute exacerbations or short term deterioration of the condition. However, rapid deterioration is rare in patients with LSS.[32]

In addition to the two subjects who walked substantially less in absolute distance during the second SPWT, one subject was identified who walked 32.9m during the first SPWT and 277.4m during the second, resulting in a 743.2% difference between tests. This individual reported using a walker at all times when walking outside the home, while the SPWT required subjects to walk without walking equipment. Anecdotal evidence suggests that this subject was uneasy about the test and not comfortable walking without the walker during the first SPWT. However, during the second SPWT this subject reported feeling more comfortable without the walker, given that the first test provided her with the confidence to do so. Therefore it is likely that the distance walked during the first test was not reflective of her true walking capacity. The results from this particular subject suggest that patients who routinely use walking equipment may need an orientation session, walking without their equipment in order to ensure that the walking test is representative of their actual capacity. Further, in future studies of walking in LSS which do not involve treadmill measurements, it may be possible to allow patients to use their walking equipment during the test. However, this subject may be an isolated case, given that similarly large percent

differences between SPWTs were not observed for the other four subjects who reported routine use of walking equipment.

5.4.3 *Distance to First Symptoms*

Distance or time to first symptoms is often reported in treadmill testing of walking capacity in LSS.[1;12;13;22;33] Results of the present study indicate that distance to first symptoms is fairly reproducible, with an ICC of 0.80 and a mean difference between testing occasions of 36.3m. However, the MAD from the present study of 214.9m is higher than expected. The only other study investigating reliability of walking time or distance to first symptoms in LSS patients was conducted by Deen et al. using a treadmill test.[1] This study reported a Concordance Correlation Coefficient (CCC) of 0.98 for time to first symptoms in the self-selected speed trial over a test re-test interval ranging from 1-4 days.[1]

5.4.4 *Walking Speed*

Although the primary construct of interest in the present study was total distance walked in the SPWT, it is also important to examine reproducibility of walking speed. It was thought that allowing patients to walk at their own pace would be most representative of everyday walking conditions. It is possible that forcing patients to walk at a speed that is different from their habitual walking speed could result in distances not representative of their true capacity. Results of the present study indicate that self-selected walking speed is highly reproducible in patients with LSS, with an ICC of 0.95, a mean difference between tests of -0.04 km and a MAD of 0.25 km/h (4% of the total range in speed). This suggests that if it were necessary to introduce a walking test time limit, permitting patients to self-select speed would still allow for some degree of standardization within patients, while maintaining the real life aspect of walking at their own pace. The high reproducibility of the walking speed also indicates that self-selected speed could be used to monitor changes in the condition.

5.4.5 *Physical Function Scale*

The results of the present study indicated that the Physical Function Scale is acceptably reproducible for the measurement of walking capacity in patients with LSS. The ICC of 0.77 was slightly lower than the expected value of 0.80.

However, ICC values of 0.75 and above are considered to be acceptable in the literature.[34] The MAD between administrations of the Physical Function Scale was 0.37, which represents 10.9% of the total range of Physical Function Scale scores. This value is less than the minimal clinically important difference of 0.52 suggested in the literature for the Physical Function Scale.[23] Reproducibility of the Physical Function Scale has been established previously in a number of studies with LSS patients. Pratt et al. reported an ICC of 0.82 over a test re-test interval of one week in a study with 29 patients with LSS and neurogenic claudication.[26] A recent study by Thornes et al. examining the Norwegian adaptation of the Swiss Spinal Stenosis Questionnaire in a sample of surgical LSS patients, reported an ICC of 0.89 over a one week interval for the Physical Function Scale.[24] Lastly, Stucki et al. reported a high Spearman correlation coefficient of 0.94 between scores for test re-test administrations of the Physical Function Scale over a period of two weeks.[10] However, the use of correlation to examine reproducibility has been questioned, given that correlation measures the strength of the linear relationship between the two variables, not agreement between them.[35] The mean Physical Function Scale score from the present study of 2.2 suggests slightly less severity on average than that found in previous studies examining the reproducibility of the Physical Function Scale, which were all conducted with subjects who were undergoing surgery for LSS.[36;37]

5.4.6 *Oswestry Disability Index*

The ODI demonstrated a high degree of reproducibility in the present study, with an ICC of 0.90, a mean difference between tests of 0.81% and a MAD of 4.1% (5.4% of the total range of ODI scores). Similar high reproducibility has been reported for the ODI in studies with low back pain populations.[27;28] However, only one study examined reproducibility of the ODI in an LSS population. This study, by Pratt et al., reported an ICC of 0.89, which is very similar to that obtained in the present study.[26] The mean ODI score in the present study is within the range of scores previously reported for studies of LSS populations.[14;38;39] Results suggest that the ODI is reproducible in LSS patients,

although the ODI was not designed specifically for use with LSS patients, nor is it intended to specifically evaluate walking capacity.

5.4.7 *Construct Specific Distance Question and Walking Distance Specific Items*

Results for the construct specific self-report question regarding predicted walking distance were not encouraging, with an ICC of 0.50 and a MAD of 837.6m (9% of the total range of self-reported distances). Of the walking measures examined, this question demonstrated the lowest reproducibility. This suggests that patients may be unable to reliably estimate their own walking capacities. These results bring into question the use of single item questions of this nature in the assessment of walking capacity in LSS. However, due to the practical issue of time constraints, the use of a single self-report question will likely persist in practice for the assessment walking capacity in this population. The walking distance questions which were isolated from the Physical Function Scale and the ODI, both appear to be more reproducible than the construct specific question discussed here. The walking distance item from the Physical Function Scale had an ICC of 0.74, a mean difference between tests of 0.04 and a MAD of 0.32. The walking distance specific item from the ODI demonstrated better reproducibility with an ICC of 0.89, a mean difference between tests of 0.14% and a MAD of 0.39%. It appears that this item from the ODI is the most reproducible of the self-report distance questions examined in the present research.

5.4.8 *Can we Generalize Results of SPWT Reproducibility to all LSS Patients?*

Total walking distance in the SPWT ranged from 32 m to 3077 m (mean 1142 m), indicating a wide variety of walking capacities and condition severities in the sample. This suggests that results of the present study could be applied to LSS patients of varying degrees of walking limitations. This notion is supported by the results of post-hoc analyses which found similar results for a subgroup of subjects who walked for the full 30 minutes and a group of subjects who did not.

In order to eliminate the possibility of a ceiling effect, the 8 subjects who walked for the full 30 minutes in the SPWT were removed and all analyses were repeated. All 8 subjects who walked for 30 minutes walked >2000m, while the

remaining 20 subjects walked <2000m. It is likely that the group of eight subjects who walked for the full 30 minutes (>2000m) represent milder severity of walking limitation, while the 20 subjects who walked <2000m represent a range from mild to severe LSS and walking limitations. Results of post-hoc analyses were similar for both groups, suggesting that the measures investigated are reproducible in LSS patients over a wide range of walking limitations. These results also eliminate the concern regarding a ceiling effect. However, larger sample sizes for each group would be required to draw any strong conclusions regarding reproducibility of walking tests for specific LSS subgroups, based on condition severity or walking limitation.

The mean distance for the eight subjects who walked for 30 minutes was 2190m±616m. Although all of these subjects had LSS and reported symptoms neurogenic claudication while walking, they appear to have a lesser degree of walking limitation and likely represent a milder condition severity compared to those whose distances were much less. No other studies of which we are aware have reported values for walking distance in LSS populations as high as this. However, this could be because most studies in the literature reporting walking distances in LSS patients involve subjects who were having surgery,[4;12;13;18;20;40] and thus could be expected to have more severe limitations. In addition, all treadmill studies in the literature reporting on walking distance imposed time limits of only 15 or 20 minutes,[4;12;13;18;20;40] as opposed to the 30 minute limit in the present study. It is possible that given a time limit of 30 minutes, subjects in previous studies may have walked further.

The mean walking distance for the subgroup of 20 subjects who did not walk for 30 minutes was 586m±416m. This result indicates that subjects in this subgroup likely have more severe LSS and walking limitations than the group of subjects who walked >2000m. Results obtained in the SPWT from this subgroup are comparable to those reported in other LSS studies with both surgical and non-surgical patients. Johnsson et al.[41] reported a slightly higher mean walking distance in 32 untreated LSS patients (630m±1063m), which was unchanged over four years, as judged by self-report. In an examination of pre-operative walking

capacity measured on a treadmill, Deen et al. reported a mean distance of 222m, [13] which is lower than that reported in the present study. Similarly, Adamova et al.[16] reported a mean distance of 236m in LSS patients with neurogenic claudication, also measured using a treadmill test.

Overall, the mean distance for the whole sample of 1142m is similar to that reported in a longitudinal study by Amundsen et al, who used self-reported walking distance in examining 100 LSS patients of varying severities.[42] This study reported that after 3 months, mean self-reported walking distance had increased from baseline values of 300m in a surgical treatment group, and 500m in a conservative treatment group to 1000m overall.[42] The reported walking distance for both groups remained stable thereafter for 10 years of follow-up.

5.4.9 *Change Scores*

For the sample as a whole there was no relationship observed between self-reported change in walking capacity and the actual difference in walking distance between SPWTs. In addition, no relationship was observed between the walking capacity change score and changes in the Physical Function Scale, ODI, self-predicted walking distance, worst pre-walking pain (measured on a VAS), or the walking distance specific items from the Physical Function Scale and ODI. These results were surprising and suggest that patients are not able to accurately perceive changes in their walking capacity over a period of up to three weeks. However, the subjects in this study may not have been able to perceive a change in their walking capacity because they do not have an accurate baseline by which to gauge change. By this we mean that on a day to day basis it is unlikely that people with LSS and walking limitations actually push themselves to their maximum walking capacity, as in the SPWT. Thus, they may not be able to perceive a change in maximum walking distance if they have not walked to capacity prior to testing, and therefore do not have a reference.

An exception to this apparent lack of relationship between the change score and actual change is the two subjects identified in this study who walked substantially less on the second SPWT. They subjects did report that their walking was "a little bit worse" before their second SPWT. It is possible that those

patients who actually have a significant and detectable change in their condition may be able to predict a change in walking capacity. Further research is warranted to examine potential factors influencing reporting of change in walking capacity, such as pain and overall health.

5.4.10 *Limitations*

A larger sample size may have provided narrower confidence intervals for reproducibility estimates and strengthened the conclusions of the present study, especially for the additional analyses of subgroups. Further research is needed to examine reproducibility of measures of walking capacity in subgroups of LSS patients, with attention given to the severity of LSS and walking limitations. Lastly, due to the study design, we were unable to obtain repeat measures of the treadmill protocol. However, reproducibility of treadmill testing in patients with LSS has been demonstrated previously.[1]

5.5 Conclusions

The results of the present study provide initial evidence regarding the reproducibility of the Self-Paced Walking Test, and confirm the reproducibility of the Physical Function Scale and the ODI in the measurement of walking capacity in LSS. This study also provides evidence to support the use of the SPWT as a criterion measure of walking capacity.

Table 5-1. Individual Differences for Repeat Administrations: SPWT Total Walking Distance

Distance SPWT 1 (m)	Distance SPWT 2 (m)	Mean Distance over Two SPWTs (m)	Absolute Difference between Tests (SPWT 1- SPWT 2) (m)	Difference as a Percent (%) of Distance in SPWT 1
43.9	43.9	43.9	0.0	0.0
721.5	722.4	722.0	0.9	0.1
2593.9	2601.5	2597.7	7.6	0.3
723.0	717.5	720.3	5.5	0.8
2487.0	2530.0	2508.5	43.0	1.7
301.8	212.8	257.3	89.1	3.0
2489.0	2590.8	2539.9	101.8	4.1
2251.6	2352.0	2301.8	100.4	4.5
965.0	911.4	938.2	53.6	5.6
2603.6	2783.4	2693.5	179.8	6.9
2473.5	2300.0	2386.8	173.5	7.0
3077.0	2844.0	2960.5	233.0	7.6
2366.2	2143.4	2254.8	222.8	9.4
153.3	167.9	160.6	14.6	9.5
1161.9	1300.0	1231.0	138.1	11.9
722.7	628.8	675.8	93.9	13.0
31.7	37.2	34.5	5.5	17.3
76.8	60.0	68.4	16.8	21.9
552.6	676.0	614.3	123.4	22.3
1628.2	1188.1	1408.2	440.1	27.0
724.5	489.8	607.2	234.7	32.4
795.8	536.4	666.1	259.4	32.6
718.7	955.2	837.0	236.5	32.9
874.8	574.2	724.5	300.6	34.3
1478.6	933.9	1206.3	533.7	36.1
102.4	142.3	122.4	39.9	39.0
368.2	706.8	537.5	338.6	92.0
32.9	277.4	155.2	244.5	743.2

Grey shading indicates subjects who walked further during the first SPWT than the second SPWT.

Table 5-2. The Difference between Test Re-test Administrations of Walking Capacity Measures in Patients with LSS (n=28)

Walking Capacity Measure	Occasion 1 (Mean±SD)	Occasion 2 (Mean±SD)	Min	Max	Mean Difference Between Occasions (Mean±SD)	Mean Absolute Deviation (MAD)*
Self-Paced Walking Test (SPWT) total distance (m)	1161.4 ±982.7	1122.4 ±963.2	31.7	3077.0	39.0±205.1	151.1
Distance to first symptoms in the SPWT (m)	480.2±737.8	516.5±716.8	8.3	2603.0	36.3±459.6	214.9
Walking speed in the SPWT (km/h)	4.1±1.1	4.1±1.1	1.5	6.2	-0.04±0.33	0.25
Physical Function Scale (1-4)	2.2±0.6	2.3±0.7	1.0	3.4	0.12±0.45	0.37
Physical Function Scale Walking Distance Specific Item (1-4)	2.2±0.7	2.3±0.8	1.0	4.0	0.04±0.6	0.32
ODI (%)	47.5±12.6	46.7±13.0	20.8	73.6	0.81±5.7	4.1
ODI Walking Distance Specific Item (1-6)	3.1±1.4	3.0±1.4	1.0	5.0	0.14±0.7	0.39
Self-predicted total walking distance (m)	1364.9±1789.6	1532.5±2036.5	25.0	9600.0	-167.6±1908.8	837.6

* MAD is calculated as the sum of the absolute differences between testing occasions/n.

Table 5-3. ICC and Confidence Intervals (CI) for Test Re-test Reproducibility of Measures of Walking in LSS (n=28)

Walking Capacity Measure	ICC Test Re-test	95% CI
Self-Paced Walking Test (SPWT) total distance (m)	0.98	0.95-0.99
Distance to first symptoms in the SPWT (m)	0.80	0.61-0.90
Walking speed in the SPWT (km/h)	0.95	0.90-0.98
Physical Function Scale	0.77	0.56-0.89
Physical Function Scale Walking Distance Specific Item	0.74	0.51-0.87
ODI	0.90	0.76-0.95
ODI Walking Distance Specific Item	0.89	0.78-0.95
Self-predicted total walking distance (m)	0.50	0.17-0.74

Table 5-4. The Difference between Test Re-test Administrations of Walking Capacity Measures in Patients with LSS Who Reported no Change between Tests (n=15)

Walking Capacity Measure	Mean Difference Between Occasions (Mean±SD)	Mean Absolute Deviation (MAD)*	Range of Individual Absolute Differences between Occasions (Time 1-Time 2)	Range of Individual Differences between Occassions Expressed as a Percentage (%) (Difference/Time1)
Self-Paced Walking Test (SPWT) total distance (m)	24.8±222.5	153.5	-544.7 – 338.6	-36.1% - 92.0%
Distance to first symptoms in the SPWT (m)	-177.2±493.8	228.4	-223.1 – 1894.0	-48.9% - 96.0%
Walking speed in the SPWT (km/h)	-0.13±0.33	0.24	-0.65 – 0.90	-22.6% - 16.0%
Physical Function Scale (1-4)	-0.17±0.51	0.44	-1.0 – 0.8	- 33.0% - 40%
Physical Function Scale Walking Distance Specific Item (1-4)	0.07±0.46	0.20	-1.0 – 1.0	-33.0% - 50%
ODI (%)	0.02±0.05	3.9	-11.32 – 5.66	-27.0% - 13.6%
ODI Walking Distance Specific Item (1-6)	0.07±0.59	0.40	-1.0 – 1.0	-25.0% - 50%
Self-predicted total walking distance (m)	31.5±607.8	348.5	-1400.0 – 1000.0	-28.0% - 83.3%

* MAD is calculated as the sum of the absolute differences between testing occasions/n.

5.6 References

1. Deen HG, Jr., Zimmerman RS, Lyons MK et al. Test-retest reproducibility of the exercise treadmill examination in lumbar spinal stenosis. *Mayo Clin.Proc.* 2000;75:1002-7.

2. Fu YS, Zeng BF, Xu JG. Long-term outcomes of two different decompressive techniques for lumbar spinal stenosis. *Spine.* 2008;33:514-8.

3. Willen J, Wessberg PJ, Danielsson B. Surgical results in hidden lumbar spinal stenosis detected by axial loaded computed tomography and magnetic resonance imaging: an outcome study. *Spine.* 2008;33:E109-E115.

4. Geisser ME, Haig AJ, Tong HC et al. Spinal canal size and clinical symptoms among persons diagnosed with lumbar spinal stenosis. *Clin.J.Pain.* 2007;23:780-5.

5. Brussee P, Hauth J, Donk RD et al. Self-rated evaluation of outcome of the implantation of interspinous process distraction (X-Stop) for neurogenic claudication. *Eur Spine J.* 2008;17:200-3.

6. Shabat S, Arinzon Z, Folman Y et al. Long-term outcome of decompressive surgery for lumbar spinal stenosis in octogenarians. *Eur Spine J.* 2008;17:193-8.

7. Weinstein JN, Tosteson TD, Lurie JD et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. *N.Engl.J Med.* 2008;358:794-810.

8. Barz T, Melloh M, Staub L et al. The diagnostic value of a treadmill test in predicting lumbar spinal stenosis. *Eur Spine J.* 2008;..

9. Katz JN, Stucki G, Lipson SJ et al. Predictors of surgical outcome in degenerative lumbar spinal stenosis. *Spine.* 1999;24:2229-33.

10. Stucki G, Daltroy L, Liang MH et al. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine.* 1996;21:796-803.

11. Fairbank JC, Couper J, Davies JB et al. The Oswestry low back pain disability questionnaire. *Physiotherapy.* 1980;66:271-3.

12. Deen HG, Zimmerman RS, Lyons MK et al. Use of the exercise treadmill to measure baseline functional status and surgical outcome in patients with severe lumbar spinal stenosis. *Spine.* 1998;23:244-8.

13. Deen HG, Jr., Zimmerman RS, Lyons MK et al. Measurement of exercise tolerance on the treadmill in patients with symptomatic lumbar spinal stenosis: a useful indicator of functional status and surgical outcome. *J Neurosurg.* 1995;83:27-30.

14. Tenhula J, Lenke LG, Bridwell KH et al. Prospective functional evaluation of the surgical treatment of neurogenic claudication in patients with lumbar spinal stenosis. *J.Spinal Disord.* 2000;13:276-82.

15. Fritz JM, Erhard RE, Delitto A et al. Preliminary results of the use of a two-stage treadmill test as a clinical diagnostic tool in the differential diagnosis of lumbar spinal stenosis. *J Spinal Disord.* 1997;10:410-6.

16. Adamova B, Vohanka S, Dusek L. Differential diagnostics in patients with mild lumbar spinal stenosis: the contributions and limits of various tests. *Eur Spine J.* 2003;12:190-6.

17. Whitehurst M, Brown LE, Eidelson SG et al. Functional mobility performance in an elderly population with lumbar spinal stenosis. *Arch Phys.Med Rehabil.* 2001;82:464-7.

18. Herno A, Airaksinen O, Saari T et al. Computed tomography findings 4 years after surgical management of lumbar spinal stenosis. No correlation with clinical outcome. *Spine.* 1999;24:2234-9.

19. Herno A, Partanen K, Talaslahti T et al. Long-term clinical and magnetic resonance imaging follow-up assessment of patients with lumbar spinal stenosis after laminectomy. *Spine.* 1999;24:1533-7.

20. Herno A, Saari T, Suomalainen O et al. The degree of decompressive relief and its relation to clinical outcome in patients undergoing surgery for lumbar spinal stenosis. *Spine.* 1999;24:1010-4.

21. Jensen OH, Schmidt-Olsen S. A new functional test in the diagnostic evaluation of neurogenic intermittent claudication. *Clin.Rheumatol.* 1989;8:363-7.

22. Moon ES, Kim HS, Park JO et al. Comparison of the predictive value of myelography, computed tomography and MRI on the treadmill test in lumbar spinal stenosis. *Yonsei Med J.* 2005;46:806-11.

23. Stucki G, Liang MH, Fossel AH et al. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J.Clin.Epidemiol.* 1995;48:1369-78.

24. Thornes E, Grotle M. Cross-cultural adaptation of the Norwegian version of the spinal stenosis measure. *Eur Spine J.* 2008;..

25. de Vet HC, Terwee CB, Knol DL et al. When to use agreement versus reliability measures. *J Clin.Epidemiol.* 2006;59:1033-9.

26. Pratt RK, Fairbank JC, Virr A. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. *Spine.* 2002;27:84-91.

27. Gronblad M, Hupli M, Wennerstrand P et al. Intercorrelation and test-retest reliability of the Pain Disability Index (PDI) and the Oswestry Disability Questionnaire (ODQ) and their correlation with pain intensity in low back pain patients. *Clin.J.Pain.* 1993;9:189-95.

28. Yakut E, Duger T, Oksuz C et al. Validation of the Turkish version of the Oswestry Disability Index for patients with low back pain. *Spine.* 2004;29:581-5.

29. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clin.Rehabil.* 1998;12:187-99.

30. Amundsen T, Weber H, Lilleas F et al. Lumbar spinal stenosis. Clinical and radiologic features. *Spine.* 1995;20:1178-86.

31. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic.Dis.* 1987;40:171-8.

32. Watters WC, III, Baisden J, Gilbert TJ et al. Degenerative lumbar spinal stenosis: an evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis. *Spine J.* 2008;8:305-10.

33. Tenhula J, Lenke LG, Bridwell KH et al. Prospective functional evaluation of the surgical treatment of neurogenic claudication in patients with lumbar spinal stenosis. *J Spinal Disord.* 2000;13:276-82.

34. Andresen EM. Criteria for assessing the tools of disability outcomes research. *Arch Phys.Med Rehabil.* 2000;81:S15-S20.

35. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1:307-10.

36. Lee J, Hida K, Seki T et al. An interspinous process distractor (X STOP) for lumbar spinal stenosis in elderly patients: preliminary experiences in 10 consecutive cases. *J.Spinal Disord.Tech.* 2004;17:72-7.

37. Sinikallio S, Aalto T, Airaksinen O et al. Depression and associated factors in patients with lumbar spinal stenosis. *Disabil.Rehabil.* 2006;28:415-22.

38. Airaksinen O, Herno A, Turunen V et al. Surgical outcome of 438 patients treated surgically for lumbar spinal stenosis. *Spine.* 1997;22:2278-82.

39. Herno A, Airaksinen O, Saari T et al. Lumbar spinal stenosis: a matched-pair study of operated and non-operated patients. *Br.J.Neurosurg.* 1996;10:461-5.

40. Herno A, Airaksinen O, Saari T. Computed tomography after laminectomy for lumbar spinal stenosis. Patients' pain patterns, walking capacity, and subjective disability had no correlation with computed tomography findings. *Spine.* 1994;19:1975-8.

41. Johnsson KE, Rosen I, Uden A. The natural course of lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 1992;82-6.

42. Amundsen T, Weber H, Nordal HJ et al. Lumbar spinal stenosis: conservative or surgical management?: A prospective 10-year study. *Spine.* 2000;25:1424-35.

CHAPTER 6

STUDY 3

Validity of Measures of Walking Capacity in Lumbar Spinal Stenosis

6.1 Introduction

Recent advances in diagnostic imaging technology, along with the aging population, have led to a marked increase in the diagnosis of lumbar spinal stenosis (LSS) internationally.[1;2] Concomitant to this, a dramatic increase in the number of surgeries for LSS has been observed over the past few decades and will no doubt continue to increase.[3] Along with the increasing prevalence of LSS diagnosis there has been a trend toward emphasizing walking capacity as a key outcome indicator for patients undergoing both surgical and conservative treatment for LSS.[4;5] Most of the recent literature focusing on treatment for LSS includes some form of walking capacity measurement.[6-8;9] As such, there is a demand for valid outcome-based measures of walking capacity in LSS patients in both research and clinical settings.

For the purpose of this research, walking capacity was defined as the distance a person with LSS is able to walk without support on a level surface at a self-selected speed before being forced to stop due to symptoms of LSS. There are a number of measures currently used with LSS patients which address to a certain degree the construct of walking capacity as defined above. However, to date there is no 'gold standard' measure of walking capacity for this population. The instruments used most frequently in the literature for this purpose include treadmill protocols,[4;6;7;10-19] the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire [20] and the Oswestry Disability Index (ODI).[21] In addition to these instruments, many authors report using single self-report questions focusing on either walking distance or time to evaluate walking capacity.[22-26] However, limited validity evidence exists regarding the use of any of these measures in the assessment of walking capacity in LSS.

6.1.1 *Treadmill Testing of Walking Capacity in LSS*

Increasing pressure from third party payers for more objective measures of outcome in LSS has led to a desire for more observational testing of function and

walking in LSS.[27] As such, the use of treadmill protocols has become more common in LSS literature and research. Treadmill testing of patients with symptomatic LSS has been shown to be safe, inexpensive, easily administered and reproducible.[4;10;11] Treadmill tests allow quantification of walking distances, as well as provocation of important clinical symptoms of LSS such as neurogenic claudication.[6] However, the recent clinical guidelines for LSS released by the North American Spine Society (NASS) concluded that there is insufficient scientific evidence to support the use of treadmill testing as an outcome or diagnostic tool with patients with LSS.[28] The limited amount of research available examining the validity of treadmill testing in LSS comes from reports by Barz et al.[6] and Tenhula et al.[7] Barz et al. reported that treadmill walking distance correlated significantly ($p<0.05$) with dural cross sectional area on MRI ($r=0.53$), the ODI ($r=0.51$) and self-reported walking distance ($r=0.62$).[6] Tenhula et al. reported a strong correlation ($r=0.88$) between treadmill walking distance and self-reported symptoms of neurogenic claudication (pain) measured using a visual analog scale (VAS).[7] However, validity of treadmill testing in LSS has yet to be examined in comparison with a criterion measure. It is not known whether measurement of walking capacity on a treadmill validly reflects walking in a more realistic setting, such as on a solid ground. Research from other clinical populations suggests that walking distance may be under-estimated using treadmill testing versus level ground walking tests, especially in older populations.[29;30] Further research is warranted to examine the validity of treadmill protocols for the measurement of walking in LSS.

6.1.2 *Self-report Measures of Walking Capacity in LSS*

Currently, treatment outcomes in patients with LSS are often determined from subjective data from self-report questionnaires,[11] such as the Physical Function Scale and the Oswestry Disability Index. The five-item Physical Function Scale of the Swiss Spinal Stenosis Questionnaire was designed specifically to evaluate physical function and walking in LSS patients, and has been used frequently for this purpose.[31-41] The items in the Physical Function Scale assess distance walked and pain limited activities of daily living involving

walking. The psychometric properties of the Physical Function Scale have received limited attention in the literature. Stucki et al.[20;42] conducted the two primary studies investigating the properties of the Physical Function Scale with a sample of 193 patients undergoing surgery for LSS. These studies found the Physical Function Scale to be internally consistent, reproducible and responsive to clinical change.[20;42] Acceptable reproducibility of the Physical Function Scale was confirmed in Study 2 (Chapter 5) of the present thesis research. However, the only pieces of validity evidence currently available to support the use of the Physical Function scale in the assessment of walking capacity in LSS come from Stucki et al. and Study 1 (Chapter 4) of this thesis.[43] Stucki et al. reported correlations between the Physical Function Scale and physicians' assessment of walking capacity (r=0.47), as well as the physical dimension of the Sickness Impact Profile (SIP) (r=0.49).[20] Study 1 of this thesis provided both convergent and divergent validity evidence regarding the use of the Physical Function Scale for the measurement of walking in an LSS population. Yet, it cannot be ascertained from either study that the construct being measured using the Physical Function Scale was in fact walking capacity. Therefore, comparison with a criterion measure of walking capacity was warranted.

The Oswestry Disability Index (ODI) has been used in many studies investigating overall function in patients with LSS.[7;16;17;44-56] It has been used once to evaluate walking distance specifically by isolating the walking distance item in a study of patients undergoing surgery for LSS.[48] In terms of psychometric properties, the ODI has been validated extensively for use in evaluating function in low back pain patients.[21;57-60] It has also been shown to be reproducible in a group of patients with LSS and neurogenic claudication (ICC=0.89),[44] as well as in Study 2 (Chapter 5) of the present thesis (ICC=0.90). However, the only evidence available to support the use of the ODI in the measurement of walking capacity in LSS comes from two studies examining the relationship between treadmill walking and the ODI in patients undergoing surgery for LSS, and Study 1 of the present research. Barz et al. reported a correlation of (r=0.51) between the ODI and total distance walked during the treadmill test,[6] while Tenhula reported a

significant correlation ($p<0.05$) between post-operative ODI scores and treadmill test parameters (correlation coefficients not reported).[7] Study 1 (Chapter 4) of the present research found the ODI to be significantly correlated with the Physical Function Scale ($r=0.72$). However, the aforementioned research does not allow us to draw any conclusions regarding the use of the ODI for measurement of walking in LSS patients. As with the Physical Function Scale, comparison with a criterion measure of walking was warranted.

Despite some validity evidence available to support the use of treadmill testing, the Physical Function Scale and the ODI in the evaluation of walking capacity in LSS, further validity investigation was warranted. It is not certain that the construct being tapped using any of these instruments or tests is, in fact, walking capacity. In order to determine whether walking capacity is the construct being measured, comparison with a criterion measure was required.

Therefore, the purpose of the present study was to examine construct validity of the various measures of walking capacity in LSS. Construct validity is the extent to which a construct behaves in accordance with theory-based hypotheses concerning how it should behave.[61] Hypotheses regarding the expected magnitude and pattern of relationships between constructs are based on theories, and the testing of these hypotheses provides validity evidence. The examination of construct validity was accomplished in the present study by forming and testing hypotheses regarding the expected relationships among measures intended to tap the construct of walking capacity (convergent), as well as measures intended to tap completely different constructs (divergent).

As part of the construct validity investigation, we examined criterion related validity evidence regarding the use of a treadmill protocol and various self-report instruments as surrogate measures of walking capacity in LSS. We were able to examine criterion validity of the surrogate measures by comparing them to a direct measure of walking capacity, otherwise know as the criterion or "gold standard".[61] The criterion in the present study was developed based on the operational definition of walking capacity in LSS described previously. The proposed criterion measure, the Self-Paced Walking Test (SPWT) asks LSS

patients to walk at their own pace, on a level surface until forced to stop due to symptoms of LSS. This test has some content related evidence supporting its construct validity given that it is a direct operationalization of the construct of interest. Although no official judgmental analysis was conducted, the relevance and representativness of the SPWT for the measurement of walking in LSS was supported by a five experts, including a spine specialist surgeon, an expert spine researcher, a methodologist and an exercise physiologist. In addition, the SPWT was found to be reproducible in Study 2 (Chapter 5) of the present thesis research.

It was hypothesized that the instruments and items intended to measure walking capacity in LSS would be highly correlated with one another and the criterion measure (>0.60), providing both criterion and convergent construct validity evidence. It was expected that the treadmill test would be the best surrogate measure of walking capacity, and mostly highly associated with the criterion.

6.2 Methods

6.2.1 *Subjects*

All subjects had been referred to one of five spine specialist surgeons in the Edmonton region with suspected LSS. At the time of their consultation with their spine specialist surgeon, subjects who met the inclusion criteria were given information about the study. The patients who expressed an interest in the study were later contacted by telephone and provided with more information. The inclusion criteria were >45 years of age with central or combination LSS confirmed on imaging (MRI/CT) and by a spine specialist surgeon. All subjects also had LSS associated walking limitations or symptoms exacerbated by walking (neurogenic claudication). Exclusion criteria included surgery for LSS within the past year, or any co-morbid conditions that would limit walking capacity or make exercise medically inadvisable as judged by the subject's physician.

6.2.2 *Data Collection*

Subjects were asked to come to the testing facility three times, with a maximum of 7 days between visits. Over the course of three visits subjects were asked to complete the Self-Paced Walking Test twice and a treadmill protocol

once. Subjects were randomized to performing either the treadmill test or the criterion SPWT during the first visit. They completed the other measure at the time of the second visit. Self-report questionnaires were completed on both occasions. Those who participated in a third visit completed the second SPWT.

Ethics approval for the study was obtained through the University of Alberta health research ethics board.

6.2.3 *Criterion Measure: Self-Paced Walking Test Protocol*

Subjects were instructed to walk continuously at their own pace around the outer lane of the track until they felt they had to stop due to symptoms of LSS (or other reasons) or until a time limit of 30 minutes had been reached. At no time were subjects encouraged or prompted to continue. Subjects were asked to indicate when they first experienced a change in symptoms, as well as indicate the nature of the symptoms (type and location). Test termination was defined as a complete stop of 3 seconds or more. Throughout the test, the administrator followed at a comfortable distance with a rolling wheel instrument to measure distance (Lufkin® Pro-Series Model PSMW38) and a stopwatch. The following information was collected: total distance and time walked, time/distance to onset of symptoms, the nature and location of symptoms (pain, numbness/tingling, weakness or fatigue), average walking speed, and the reason for test termination should they not walk for the full 30 minutes (symptoms of LSS, fatigue, shortness of breath, dizziness, pain or discomfort due to co-morbidities).

6.2.4 *Treadmill Protocol*

The treadmill test used in the present study was a modification of the protocol described by Deen et al.[4;10;11] Subjects were asked to walk on a treadmill at 0% grade and a self-selected speed until they felt they had to stop due to symptoms of LSS (or other reasons), or until a time limit of 30 minutes had been reached. To start the test, the administrator slowly increased the speed to 1.2 miles per hour. Subjects were then asked if they would like to modify the speed to find a pace that was comfortable for them. Subjects were also allowed to adjust the speed throughout the test. Speed modifications were done by the test administrator. Subjects were asked to notify the administrator when they first

experienced a change in symptoms and to indicate the nature of the symptoms. At no time were subjects encouraged or prompted to continue. Subjects were asked to avoid holding the hand rails so as to maintain an upright posture. Subjects were permitted to place their hand, palm side down under the rail if it made them more comfortable or to steady themselves. The same data were collected as in the SPWT.

6.2.5 *Questionnaire*

Subjects completed a self-report questionnaire before walking during each of the first two visits. Information on participant characteristics of age and gender was acquired, as well as history of the current condition, including presence of back and/or leg pain and duration of back and/or pain. The battery of standardized measures included the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire,[20] the Oswestry Disability Index,[21] a 100mm visual analog scale for walking difficulty adapted from Yamashita et al.,[5] and the Health Utilities Index Mark 2 and 3 (HUI 2/3), a preference-based measure addressing health related quality of life.[62] In addition to the aforementioned questionnaires, a number of items from these instruments that addressed walking distance specifically were isolated for analysis. These items included an item regarding pain limited walking distance from the ODI (#4), the HUI3 Single Attribute Utility Score for Ambulation, and the first item from the Physical Function Scale which addresses walking distance.

The Physical Function Scale score was calculated as the un-weighted mean of the five items in the scale. The resulting possible scores of 1-4 represent a range from mild to severe limitation in physical function/walking.[20] The item addressing walking distance included the following options: ability to walk "over 3 km", "over 250m but less than 3 km", "over 15m but less than 250m", or "less than 15m".[9] The ODI was calculated as a percentage of the total possible score of 53, with a greater score representing greater back related disability.[21] The item addressing pain limited walking distance from the ODI included the following options: "I am in bed most of the time", "I can only walk using a stick or crutches", "Pain prevents me from walking more than 500m", "Pain prevents me

from walking more than 1km", "Pain prevents me from walking more than 2km", "Pain does not prevent me from walking any distance". The HUI3 Ambulation and Cognition Scores are derived using specific HUI algorithms, with a score ranging from 0 (most disabled) to 1 (no disability).[62] The HUI3 Global score is calculated using 8 single attribute vectors (vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain), on a scale from 0 (dead) to 1(healthy).[62]

Subjects were also asked the following two construct specific questions: "How would you say your walking capacity is today, on a scale of 0 to 10, with 0 being your worst day, and 10 being your best day?" and "If you were to go for a walk today, how far would you be able to walk at your own pace, on level ground before being forced to stop due to symptoms of lumbar spinal stenosis (m/km)?" In order to provide the same distance reference to all subjects, they were informed that a straight-away on a track which was in view was 60m (~200ft).

Before and after each walking test subjects completed the Visual Analog Pain Scales (VAS) for back, right leg and left leg and indicated on a body diagram the location and nature of their symptoms.

If subjects were taking medications, they were asked to be consistent with their medications on the testing days. In addition, they were asked to do their best to stay consistent with any other treatments they were receiving between visits so as to minimize any changes in their condition.

6.2.6 *Hypotheses for Convergent and Divergent Construct Validity*

High convergent correlations (>0.60)[63] were expected among items intended to measure walking capacity, including the SPWT final distance, treadmill test final distance, self-reported walking distance, Physical Function Scale mean, 0-10 rating for walking capacity, visual analog scale for walking difficulty (VAS), [5] HUI3 Ambulation score, and walking distance specific items from the ODI and the Physical Function Scale. Moderate convergent correlations (>0.40) were expected between the walking measures and both the Oswestry Disability Index and the HUI3 Global Utility Score. It was also hypothesized that the walking specific measures would be divergently correlated (<0.30) with the

measures intended to tap completely different constructs, including age and the HUI3 Cognition score.

6.2.7 *Hypotheses for Criterion Validity*

The criterion variable was defined as the final distance measured in the first SPWT (m). It was hypothesized that the instruments and items intended to measure walking capacity in LSS would be convergently correlated with the final distance measured in the SPWT (>0.60) providing criterion related validity evidence. These items included the treadmill test final distance, self- reported walking distance, Physical Function Scale mean, 0-10 rating for walking capacity, visual analog scale for walking difficulty (VAS)[5], HUI3 Ambulation score, and walking distance specific items from the ODI and the Physical Function Scale. It was hypothesized also that the distance measured in the Self-Paced Walking Test would correlate highest with the distance from the treadmill test (r>0.80). A high level of agreement was expected between the SPWT and treadmill final distances.

6.2.8 *Data Analysis*

The previously stated hypotheses regarding expected relationships among measures were examined using a modification of the correlational method suggested by Campbell and Fiske [64] (multi-method correlation matrix). Spearman correlations coefficients were used to determine associations between measures, given that some of the items were measured on an ordinal scale. An alpha (α) level of 0.05 was chosen to judge significance, with an expected power of 0.80. Associations of 0.30-0.59 were judged to provide moderate convergent evidence, >0.60 strong convergent evidence, and >0.80 were very strong convergent evidence. Associations <0.30 were judged to be low and provide divergent validity evidence.[63] However, most important was the relative strength of the correlations to one another, supporting the hypothesized relationships to the construct.

A paired *t*-test was used to determine whether the mean walking distance on the SPWT was significantly different from the mean distance in the treadmill test. The level of agreement between the treadmill protocol and the SPWT was also examined using mean absolute deviation (MAD), which is the sum of the

absolute differences between tests divided by the number of subjects. If two tests are in perfect agreement there would be a MAD of zero. The MAD was employed given that it is not affected by the direction (+/-) of the difference. SPSS for Windows, version 15.0 was used for all statistical analysis (SPSS Inc., Chicago, Illinois).

In anticipation of a possible ceiling effect due to subjects walking for the full 30 minutes, subjects who walked for 30 minutes were removed post-hoc from the analysis and all analyses were repeated.

6.3 Results

6.3.1 *Subjects*

The study sample consisted of 41 subjects with a mean age of 66.9±9.6 years, 23 of whom were women. All patients reported having both back and leg pain, with a mean duration of 10±11 and 7±9 years respectively. Seven subjects reported routine use of walking equipment (3 walkers, 4 canes).

6.3.2 *Walking Tests*

Mean scores for the walking capacity measures are shown in Table 6-1. The mean self-selected speed for the SPWT (4.1±1.1km/h or 2.5±0.7mph) was significantly higher than that of the treadmill test (2.5±1.0km/h or 1.5±0.6mph) ($p<0.05$). The mean total walking time for the SPWT (912.0±665.5 seconds) was significantly higher than the total time measured on the treadmill (770.7±673.0 seconds) ($p<0.05$).

6.3.3 *Convergent and Divergent Construct Validity Evidence*

All correlations are presented in Table 6-2. Hypotheses were supported regarding the relative order and magnitude of correlations among the measures intended to tap the construct of walking capacity, and those intended to measure different constructs (age and cognition). Convergent validity evidence included correlations ($r=0.52$-0.88) among all of the measures intended to tap the construct of walking capacity (with the exception of the VAS for walking difficulty) including the SPWT final distance, treadmill test final distance, self-reported walking distance, Physical function scale mean, 0-10 rating for walking capacity, HUI3 ambulation score, and walking distance specific items from the ODI and the

Physical Function Scale. The majority of correlations among the walking measures were >0.60, as hypothesized (Table 6-2). Divergent validity evidence included low correlations (<0.30) between the measures intended to tap the construct of walking capacity and both age and the HUI3 Cognition Score. All divergent correlation coefficients were <0.30 as hypothesized. These results provide both convergent and divergent validity evidence for the use of the SPWT, the treadmill test and the self-report instruments and items in the measurement of walking capacity in LSS.

6.3.4. *Criterion Validity Evidence: Treadmill*

As hypothesized, the distance measured in the treadmill test was correlated highest with the distance measured in the SPWT (r=0.88) (Table 6-2). While this high correlation between the SPWT and treadmill test indicates a high level of agreement in the ranking of subjects, the absolute distances between the two tests were found to be significantly different ($p<0.05$). A substantial bias was observed with 37/41 (90%) subjects walking further on the SPWT than on the treadmill test. The significant mean difference between the SPWT and the treadmill test was 365m ($p<0.05$), while the mean absolute deviation (MAD) between the two tests, (which does not account for the direction of the difference) was 392.8m. Table 6-3 describes the individual absolute differences in distance between the SPWT and treadmill test, ranked from smallest to largest percent difference (difference in distance between tests as a percent of the SPWT distance). The absolute difference in distance for the four subjects who walked further during the treadmill test range from 10.2m-252.3m. The remaining 37 subjects who walked further during the SPWT can be broken down into categories based on the absolute differences between tests. Nine subjects walked less than 50m further on the SPWT (range of 10.7m-45.4m), five subjects walked between 50m-200m further (range of 75.0m-110.7m), nine subjects walked between 200m-400m further (range of 206.5m-237.4m), six walked between 400m-600m further (range of 449.1m-585.1m), five walked between 600m-100m further (range 657.8m-959m) and three subjects walked more than 1000m further on the SPWT (range 1127.0m-2106.8m). If the subjects who walked further on the SPWT are

broken down using the percent by which the treadmill under-estimated the SPWT, two walked less than 10% further on the SPWT, nine walked between 10-25% further, fourteen walked between 25-50% further, six walked between 50-75% further and six walked between 75-100% further. Overall including all subjects, the treadmill test under-estimated the distance measured on the SPWT by approximately 25%. However, the mean difference was unduly affected by one outlier who walked 329% further on the treadmill test. After the removal of this one outlier, the treadmill test under-estimated the SPWT distance by an average of 35%.

6.3.5 *Criterion Validity Evidence: Self-Report Measures*

Hypotheses regarding convergent correlations (>0.60) between the SPWT and the self-report measures of walking capacity were supported, providing criterion validity evidence for the use of these measures in the assessment of walking capacity in LSS, with the exception of the VAS for walking difficulty (r=0.41) (Table 6-2). Of the self-report measures, the individual item from the ODI addressing pain limited walking distance was most highly correlated with the criterion (r=0.82), followed by the walking distance item from the Physical Function Scale (r=0.72), the HUI3 Ambulation Score (r=0.68), the Physical Function Scale as a whole (r=0.68), walking capacity rated on a 0-10 scale (r=0.66) and self-predicted walking distance (r=0.65) (Table 6-2). When examining agreement between the SPWT and self-predicted walking distance, subjects tended to over-estimate their actual walking distance by approximately 30% (Table 6-1).

6.3.6 *Post-hoc Analyses*

The possibility of a ceiling effect was a concern, given that 10 of 41 subjects (24%) walked for the full 30 minutes. However, when these subjects were removed from the analysis, results were similar (Table 6-4). The pattern of correlation coefficients was similar for these 33 subjects, while the magnitude of some validity coefficients was slightly lower. The treadmill (r=0.81) and the ODI walking distance item (r=0.79) maintained the highest correlations with the criterion SPWT in this post-hoc analysis. However, with the 10 subjects removed,

one difference was noted. Instead of subjects over-estimating their walking capacity by 30%, subjects tended to underestimate their walking capacity by 30%. The mean walking distance for the SPWT with the 10 subjects removed was 513.6 m ± 440.8 and 362.0 m ± 463.9 for self-predicted walking distance.

6.4 Discussion

The results of the present study provide both construct and criterion validity evidence for the use of a treadmill protocol and various self-report instruments, including the Physical Function Scale and the walking distance item from the ODI in the measurement of walking capacity in LSS. The validity evidence suggests that the construct being tapped using the treadmill test, Physical Function Scale and ODI walking item was likely walking capacity. Results suggest that the treadmill protocol is the best surrogate measure of walking capacity, although it tends to substantially under-estimate distance measured in the SPWT.

6.4.1 *The SPWT as the Criterion Measure of Walking Capacity in LSS*

Some readers may question the use of the SPWT as a criterion for walking capacity, given that it is a new measure. However, it has been suggested in prominent validity literature that an unquestionable criterion may be established as a consequence of an operational definition of the construct of interest.[61] The operational definition on which the SPWT is based encompasses the aspects of walking which are most relevant to and representative of LSS patients' actual walking capacities in real life situations, including functional distance, level ground walking, self-selected speed and a symptom limited end-point. The SPWT is a direct operationalization of the construct of walking capacity in LSS as defined in this research. This measure is as relevant to and representative of the defined construct of walking capacity in LSS as possible. As previously mentioned, the relevance and representativeness of the SPWT were supported through informal judgemental analysis. In addition, The SPWT was found to be highly reproducible, as reported in Study 2 (Chapter 5), suggesting that correlations with the surrogate measures were not attenuated by a lack of reproducibility of the criterion. The only aspect of the test which may not have

been realistic for some patients is that they were unable to use supportive devices for the test. This decision was required in order to compare the criterion with the treadmill test, given that it is not possible to use walking aids on a treadmill.

6.4.2 *Validity of Treadmill Measurements*

This is the first study of which we are aware that was designed to examine validity of a treadmill test for use in measurement of walking capacity in LSS. It has been suggested in the present study that an observational level ground walking test such as the SPWT can be used the gold standard measure of walking capacity in LSS. However, the SPWT may not be a feasible option in many clinical settings. It is likely that treadmill testing may be the more viable option for observational testing of walking in LSS populations, given that most clinical and research personnel have access to a treadmill. However, while both construct and criterion validity evidence was provided in this study for the use of the treadmill test in the measurement of walking capacity in LSS, results suggest that the treadmill test substantially under-estimates the distance measured on the SPWT.

As expected, the treadmill test was found to be most highly correlated with the criterion ($r=0.88$), as well as convergently correlated (>0.60) with the self-report measures of walking capacity ($r=0.63$-0.83). These results are consistent with the limited research in this area, which has shown that treadmill walking distance correlated significantly ($p<0.05$) self-reported walking capacity ($r=0.62$)[6] and self-reported symptoms of neurogenic claudication ($r=0.88$).[7] However, despite the high correlation between the treadmill test and the SPWT, a significant bias was observed, with 37/41 patients walking further in the SPWT than on the treadmill. While the high correlation between the SPWT and treadmill tests suggests that the tests are able to rank subjects similarly, on examining absolute distances walked, there was a substantial difference between tests, with a mean difference of 365m and a MAD was 392.8m. Of the 37 subjects who walked further on the SPWT than on the treadmill the majority (63%) walked at least 25% further on the SPWT and 30% walked at least 50% further. Eleven subjects (27%) walked less than 25% further during the SPWT, and only two (5%) walked

less than 10% further. Furthermore, on examining absolute differences between the tests, the majority of subjects (56%) walked at least 200 m further on the SPWT and fourteen (34%) walked more than 400 m further. Only four subjects (10%) walked further on the treadmill test. Therefore, these results suggest that if measurement of actual walking distance is of importance, the SPWT is definitely preferable to a treadmill test, given that the treadmill significantly under-estimates walking capacity.

The difference in distance observed between the SPWT and treadmill test can be explained by patients walking significantly faster and for greater total time during the SPWT as compared to the treadmill, resulting in a systematic bias. Observed differences are likely due to the unnatural nature of treadmill walking, and the aged population of patients with LSS who may not feel comfortable on a treadmill. Anecdotal reports from patients in the present study confirmed feelings of discomfort, unsteadiness, boredom and unease on the treadmill. A number of subjects who walked much less on the treadmill reported that they had never used a treadmill before and were very uncomfortable and intimidated walking on one. The subject who had the greatest absolute difference in distance between tests (2106.8m) indicated that she did not feel balanced on the treadmill. Although all subjects who did not walk for 30 minutes did report stopping the test due to symptoms of LSS, it is possible that patients were stopping the test earlier on the treadmill due to discomfort, unease or balance issues, and not entirely due to LSS symptoms. This is a potential limitation to the use of treadmill testing with LSS patients, especially in those who are frail, have balance problems, or are used to walking with aids (e.g. cane or walker). Findings from the present study corroborate those from a study comparing treadmill and corridor walking tests in older patients with chronic obstructive pulmonary disease (COPD). Swerts and colleagues reported that both walking speed and distance were higher in a 12 minute self-paced corridor test compared to a 12 minute self-paced treadmill test. It was concluded that the differences were likely attributable to patients' increased familiarity with level ground walking.[29]

Although the results of this study suggest that a SPWT is preferable to a treadmill test, the reality is that treadmill testing is currently used in clinical settings, and will no doubt continue to be used. As previously mentioned, the other instruments being used in clinical and research practice to measure walking in LSS are self-report questionnaires. If it is not possible to conduct a SPWT, treadmill testing is a better alternative than self-report measures, given the clinical utility of observing patients' walking. Self-report questionnaires cannot capture the actual nature and timing of symptoms elicited during an observational test, nor can questionnaires really capture aspects of functional walking ability such as gait characteristics, speed and balance issues. Further, despite the obvious bias in measurement, the treadmill test was found to be most highly correlated with the criterion, as well as highly correlated with the self-report measures of walking capacity. This high correlation between the SPWT and the treadmill test indicates that the two tests rank subjects the same, and that the treadmill test is capturing the construct of walking capacity to some degree. Therefore, it may be reasonable to use a treadmill test of walking capacity with LSS populations, with the assumption that patient ranking would be similar to ranking obtained using a SPWT, yet the absolute distances measured in a treadmill test are likely under-estimate patients' true walking capacities by a substantial degree (approximately 25-35%).

6.4.3. *Validity of Self-Report Instruments and Items*

Although measurement of walking capacity using an SPWT or treadmill test is thought to be clinically preferable, the reality is that many health professionals do not have the time or resources to use such tests. Assessment of walking capacity is often accomplished using self-report instruments or items. Consequently, we examined the validity of a number of instruments and items currently in use in research and clinical practice. Construct and criterion validity evidence was provided for the use of all self-report measures and instruments used to measure walking capacity in the present study, with the exception of the walking difficulty item. The order and magnitude of correlations among the measures of walking, and the measures of different constructs (age, cognition)

were as expected, providing both convergent and divergent validity evidence for the use of these measures in assessing walking capacity in LSS. As hypothesized, all measures (except the walking difficulty item) were correlated >0.60 with the criterion, providing criterion validity evidence. The self-report item found to be most highly correlated with the criterion SPWT was the walking distance item isolated from the ODI (r=0.82).

6.4.4 *Physical Function Scale*

The mean Physical Function Scale score from the present study of 2.3 is similar to that reported in Study 1 (2.1±0.7),[43] and suggests slightly less severity on average than reported in previous studies of subjects who were all undergoing surgery for LSS.[32;65;66] Conclusions from this study are likely to apply most accurately to those patients with Physical Function scale scores in the range of 2.3±0.7.

In terms of validity, the Physical Function Scale was found to correlate convergently with the criterion SPWT (r=0.68), as well as with the other measures intended to tap the construct of walking capacity (Table 6-2). The results of this study support the conclusions of Study 1 (Chapter 4), while providing evidence that the construct being measured using the Physical Function Scale is likely walking capacity. Study 1 found the Physical Function Scale to be significantly ($p<0.05$) correlated with the ODI (r=0.72), the ODI walking distance item (r=0.80) and the HUI3 ambulation score (r=0.62), supporting convergent construct validity for the measurement of walking capacity. Results of the present study compare favourably with those reported by Thornes et al. who found the Physical Function Scale to be correlated significantly ($p<0.05$) with both the ODI (r=0.70) and the visual analog scale for leg pain (r=0.41),[67] and those reported by Stucki et al. who reported a correlation of r=0.47 with physicians' assessments of walking capacity.[20]

Although the Physical Function Scale as a whole is thought to reflect walking capacity in LSS, results of the full scale may be difficult to interpret if you are interested in walking distance specifically. Possible scores of 1-4 for the Physical Function Scale represent a range from mild to severe limitation in

physical function/walking, not specific distances.[20] As such, the distance specific item from this scale was isolated for analysis. This item was found to be correlated r=0.72 with the criterion and r=0.63 with the treadmill distance, suggesting that it may be feasible to isolate this item to evaluate walking distance specifically. This item was found to be acceptably reproducible in Study 2 (Chapter 5) of this research (ICC=0.74), although the ICC was lower than expected.

6.4.5 *Oswestry Disability Index*

Given the widespread use of the ODI in LSS and low back pain research, it was important to determine whether the ODI is valid for use in evaluating walking capacity, to allow for comparison of walking capacities between LSS studies, and between different low back conditions. The mean ODI score of 47%±13 in the present study is within the range of scores reported in the literature for LSS populations.[46;47;50] In terms of validity, the ODI as a whole was not highly correlated with the criterion SPWT (r=0.41), suggesting that the full ODI does not reflect the construct of walking capacity. However, much like the distance specific item from the Physical Function Scale, the walking distance specific item from the ODI was isolated for analysis. This single item was found to have convergent correlations (r=0.66-0.82) with other walking measures, including the criterion SPWT distance (r=0.82) and the treadmill distance (r=0.83). This item was found to be reproducible in Study 2 (ICC=0.89) (Chapter 5). The high convergent correlation with the criterion suggests that this item ranks subjects the same as the SPWT and is tapping the construct of walking capacity in LSS. Therefore, it may be possible to use the ODI as a whole to evaluate back pain related disability, allowing comparison across multiple spinal disorders, while isolating the walking item to evaluate walking capacity specifically in studies including LSS patients. However, as a caution, this item only provides distance information in the following categories: "I am in bed most of the time", "I can only walk using a stick or crutches", "Pain prevents me from walking more than 500m", "Pain prevents me from walking more than 1km", "Pain prevents me from walking more than 2 km", "Pain does not prevent me from walking any distance".

Therefore, if the construct of interest is actual walking distance, the information provided from this item is limited.

6.4.6 *Self-Reported Walking Distance*

The construct specific question that asked patients to estimate walking capacity in metres was found to be convergently correlated with both the SPWT (r=0.65) and the treadmill test (r=0.73), as well as the other measures of walking capacity, providing both construct and criterion validity evidence for its use in evaluating walking capacity in LSS. However, on examining actual agreement between the two measures, with the whole group included, subjects tended to overestimate actual walking capacity by approximately 30% (300m in mean), compared to the distance measured in the SPWT. Yet, it is likely that many of the subjects who walked for the full 30 minutes would over-estimate their walking distance as measured on the SPWT, given that they were stopped at 30 minutes, and could not reach their actual maximum distance. When these 10 subjects who walked for the full 30 minutes were removed, the remaining 31 subjects actually tended to under-estimate their walking capacity as measured on the SPWT by approximately 30%. This suggests that more limited patients may tend to under rather than over-estimate their actual capacity. These results suggest that although this construct specific item was convergently correlated with the criterion, suggesting that the two measures rank subjects the same, subjects are likely to under-estimate their actual walking capacity by approximately 30%. In addition, of the walking measures examined in Study 2 (Chapter 5), this question showed the worst reproducibility, with an ICC of 0.50.

6.4.7 *Can we Generalize Validity Results to All LSS Patients?*

Given that to date no walking tests similar to the SPWT have been used in LSS research, we cannot compare SPWT results to similar literature. In terms of the treadmill literature, the mean treadmill distance from the present study (630m±683m) was very similar to that reported by Johnsson et al. in a treadmill study of 32 untreated patients with LSS (630m±1063m).[68] However, mean treadmill distances from the present study are higher than values reported in the literature for LSS patients undergoing surgery, suggesting that the present sample

may represent patients who are less limited. Deen et al. reported a mean distance of 222m in an examination of pre-operative walking capacity measured on a treadmill,[11] and Adamova et al. reported a mean distance of 236m in LSS patients with neurogenic claudication, also measured using a treadmill test.[13] While the mean values for the SPWT and treadmill test in the present study are higher than many values reported in the literature, the wide variation in walking distances observed in both the SPWT (10-3077m) and treadmill test (0-2205m) indicates that this sample includes patients of a wide range of severities of walking limitation. If we break the subjects down into distance categories for the SPWT as recommended by Ogibuko,[69] eight (20%) walked less than 100m (very limited), seven (17%) walked more than 100m but less than 500m (moderately limited), fourteen (34%) walked more than 500m but less than 1000m (limited) and twelve (29%) walked more1000m (minimally limited). It is apparent that the current sample includes patients of all severities, suggesting that results can be generalized across LSS patients with varying walking limitations.

6.4.8 *Limitations*

The primary limitation of the present study was subject recruitment and sample size. However, the magnitude and significance of observed correlations coefficients were as expected, and sufficiently powered. The possibility of a ceiling effect was a concern, given that 10 of 41 subjects (24%) walked for the full 30 minutes. However, when these subjects were removed from the analysis, results were similar.

We did consider determining a bias correction factor for the treadmill to allow for use of the treadmill test in place of a SPWT. However, the small sample size and heterogeneity of the differences in distance between the treadmill and SPWT precluded the determination of such a correction factor. It may be possible to determine a correction factor for treadmill tests of walking in LSS in future research using a much larger sample size. Further, different correction factors for subgroups of a more homogeneous nature, based on differences in distance walked, may be needed.

Overall, the results of this research do suggest that in terms of observational tests of walking, the use of a self-paced, level ground walking test such as the SPWT is preferable to a treadmill test. However, as mentioned previously, tests such as the SPWT are not always feasible. It would be worthwhile in future research to explore other options for measurement of walking in LSS. It is possible that instead of using a treadmill test as an alternative to the SPWT, the use of community based walking tests (administered by patients themselves) or pedometers may allow for the collection of data similar to that obtained in an SPWT, while eliminating the concern of time burden for health professionals. If patients could measure their own walking capacity in a realistic setting, this would reduce the need for testing in clinics or hospitals. Further research is warranted to examine the feasibility of such options.

There is one factor yet to be discussed that may affect the validity of the SPWT. Although the SPWT is thought to be a symptom-limited test, we can never know for sure that a person actually stopped the test due to symptoms of LSS. There are a number of other potential factors which could influence a persons' decision to stop. These include, but are not limited to motivation, interaction with the tester (although this was minimized during the actual test), non-LSS related pain or discomfort and fear of a maximal effort. However, all subjects who did not reach the 30-minute time limit reported stopping due to symptoms of LSS.

6.5 Conclusions

The results of the present study have implications for both research and clinical settings involving patients with LSS. Both construct and criterion validity evidence was provided regarding the use of a treadmill protocol, the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire and various other self-report instruments in the measurement of walking in LSS. Although a strong relationship was demonstrated between the treadmill protocol and the criterion SPWT, a systematic bias was observed with patients walking significantly further in the SPWT. It is recommended that if possible, an observational level ground walking test similar to the SPWT be used in evaluating walking capacity in LSS,

when measurement accuracy is of importance. However, if such a test is not feasible, validity evidence has been provided for the use of a treadmill protocol and various self-report instruments for this purpose.

Table 6-1. Walking Capacity Measures (n=41)

Walking Measure	Mean±SD
Total distance walked in Self-Paced Walking Test (m)	994.4±945.6
Total time in the Self-Paced Walking Test (seconds)	912.0±665.5
Total distance walked in treadmill test (m)	630.1±683.6
Total time in the treadmill test (seconds)	770.7±673.0
Physical Function Scale	2.36±0.7
ODI (as a percentage)	47.0±12.9
Self-rated walking capacity (0-10)	5.4±2.2
Self-reported predicted walking distance (m)	1292.7±2033.8
VAS for walking difficulty (0-10)	5.5±2.3
HUI3 Ambulation Score (0-1)	0.8±0.2

Table 6-2. Convergent and Divergent Validity Coefficients among Measures of Walking Capacity in Lumbar Spinal Stenosis (n=41)

Spearman's Rho (n=41)	SPWT	Tread	PF walk	ODI walk	HUI3 amb	PF Scale	Walk cap	Walk dist	Walk diff	ODI	HUI3 Glob	Age	HUI3 Cog
SPWT	1.00												
Tread	0.88**	1.00											
PF walk	0.72**	0.63**	1.00										
ODI walk	0.82**	0.83**	0.76**	1.00									
HUI3 Amb	0.68**	0.66**	0.68**	0.68**	1.00								
PF Scale	0.68**	0.67**	0.72**	0.80**	0.56**	1.00							
Walk Cap	0.66**	0.60**	0.64**	0.66**	0.64**	0.65**	1.00						
Walk dist	0.65**	0.73**	0.58**	0.71**	0.52**	0.57**	0.62**	1.00					
Walk diff	0.44*	0.52*	0.43**	0.56**	0.42**	0.62**	0.56**	0.32*	1.00				
ODI	0.41*	0.40*	0.39*	0.60**	0.41**	0.65**	0.47**	0.39*	0.52**	1.00			
HUI3 Glob	0.35*	0.47**	0.40*	0.40*	0.53**	0.63**	0.37*	0.26*	0.57**	0.72**	1.00		
Age	0.02	0.13	0.09	0.14	0.03	0.02	0.03	0.12	0.16	0.12	0.07	1.00	
HUI3 Cog	0.09	0.23	0.05	0.20	0.04	0.14	0.06	0.04	0.19	0.03	0.24	0.40*	1.00

** p<0.01, *p<0.05. Gray shading indicates divergent correlation coefficients, while no shading indicates convergent coefficients. SPWT = Self-Paced Walking Test Final Distance (m); Tread = Treadmill Test Final Distance (m); PF walk = Walking Distance Item from the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire; ODI walk = Walking Distance Item from the Oswestry Disability Index; HUI3 Amb = Single Attribute Utility Score for Ambulation; PF Scale = Physical Function Scale of the Swiss Spinal Stenosis Questionnaire; Walk Cap = Walking Capacity on a 0-10 Scale from Worst to Best; Walk dist = Self Predicted Walking Distance (m); Walk diff = Visual Analog Scale for Walking Difficulty; ODI = Oswestry Disability Index; HUI3 GLOBAL = Health Utilities Index Mark 3 Global Health Related Quality of Life Utility Score; HUI3 Cog = Single Attribute Utility Score for Cognition.

Table 6-3. Individual Differences in Distance between Self-Paced Walking Test and Treadmill Test

SPWT Distance (m)	Treadmill Distance (m)	Absolute Difference between Tests (SPWT –Treadmill Distance) (m)	Difference as a Percent (%) of SPWT Distance
715.7	675.9	39.8	5.6
1628.2	1738.0	109.8	6.7
102.4	112.7	10.2	10.0
557.8	482.8	75.0	13.4
281.6	241.4	40.2	14.3
2603.6	2204.8	398.8	15.3
2473.5	2092.0	381.5	15.4
2206.8	1824.0	382.8	17.3
552.6	450.6	102.0	18.5
61.6	48.3	13.3	21.6
2489	1931.0	558.0	22.4
1161.9	885.1	276.8	23.8
43.9	32.2	11.7	26.7
721.5	515.0	206.5	28.6
724.5	515.0	209.5	28.9
368.2	257.5	110.7	30.1
2593.9	1770.3	823.6	31.8
2487.0	1528	959.0	38.6
118.0	80.5	37.5	38.8
2301.0	1400.1	900.9	39.2
1478.6	820.8	657.8	44.5
723.0	386.2	336.8	46.6
2366.2	1239.2	1127.0	47.6
965.0	502.2	462.8	48.0
31.7	16.1	15.6	49.2
3077.0	1561.0	1516.0	49.3
718.7	338.0	380.7	53.0
76.8	32.2	44.6	58.1
722.7	273.6	449.1	62.1
795.8	267	528.8	66.5
874.8	289.7	585.1	66.9
153.3	48.3	105.0	68.5
360.0	612.3	252.3	70.1
301.8	64.4	237.4	78.7
56.7	11.3	45.4	80.1
944.0	88.5	855.5	90.6
600.0	48.3	551.7	92.0
2251.6	144.8	2106.8	93.6
10.7	0.0	10.7	100.0
32.9	96.6	63.7	193.6
37.5	160.9	123.4	329.3

Grey shading indicates subjects who walked further during the treadmill test than the SPWT.

Table 6-4. Convergent and Divergent Validity Coefficients among Measures of Walking Capacity in Lumbar Spinal Stenosis Excluding Subjects who Walked a Full 30 Minutes in the SPWT (n=33)

Spearman's Rho (n=33)	SPWT	Tread	PF walk	ODI walk	HUI3 amb	PF Scale	Walk cap	Walk dist	Walk diff	ODI	HUI3 Glob	Age	HUI3 Cog
SPWT	1.00												
Tread	0.81**	1.00											
PF walk	0.51**	0.51**	1.00										
ODI walk	0.79**	0.67**	0.65*	1.00									
HUI3 Amb	0.60**	0.55**	0.58*	0.51**	1.00								
PF Scale	0.66**	0.52**	0.65**	0.72**	0.40**	1.00							
Walk Cap	0.55**	0.35*	0.39*	0.52**	0.54**	0.54**	1.00						
Walk dist	0.58**	0.76**	0.30*	0.64**	0.48**	0.50**	0.44**	1.00					
Walk diff	0.30*	0.10	0.18	0.24	0.26	0.37*	0.44**	0.04	1.00				
ODI	0.53*	0.54*	0.40*	0.62**	0.31	0.68**	0.50**	0.42*	0.57**	1.00			
HUI3 Glob	0.55*	0.49**	0.47*	0.40*	0.54**	0.63**	0.33*	0.26*	0.54**	0.72**	1.00		
Age	0.10	0.03	0.09	0.04	0.12	0.16	0.09	0.07	0.02	0.18	0.24	1.00	
HUI3 Cog	0.01	0.03	0.12	0.03	0.08	0.05	0.08	0.02	0.34	0.07	0.11	0.37*	1.00

** $p<0.01$, *$p<0.05$. Gray shading indicates divergent correlation coefficients, while no shading indicates convergent coefficients. SPWT = Self-Paced Walking Test Final Distance (m); Tread = Treadmill Test Final Distance (m); PF walk = Walking Distance Item from the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire; ODI walk = Walking Distance Item from the Oswestry Disability Index; HUI3 Amb = Single Attribute Utility Score for Ambulation; PF Scale = Physical Function Scale of the Swiss Spinal Stenosis Questionnaire; Walk Cap = Walking Capacity on a 0-10 Scale from Worst to Best; Walk dist = Self Predicted Walking Distance (m); Walk diff = Visual Analog Scale for Walking Difficulty; ODI = Oswestry Disability Index; HUI3 GLOBAL = Health Utilities Index Mark 3 Global Health Related Quality of Life Utility Score; HUI3 Cog = Single Attribute Utility Score for Cognition.

6.6 References

1. Kent DL, Haynor DR, Larson EB et al. Diagnosis of lumbar spinal stenosis in adults: a metaanalysis of the accuracy of CT, MR, and myelography. *AJR Am J Roentgenol.* 1992;158:1135-44.

2. Lurie JD, Birkmeyer NJ, Weinstein JN. Rates of advanced spinal imaging and spine surgery. *Spine.* 2003;28:616-20.

3. Deyo RA, Gray DT, Kreuter W et al. United States trends in lumbar fusion surgery for degenerative conditions. *Spine.* 2005;30:1441-5.

4. Deen HG, Jr., Zimmerman RS, Lyons MK et al. Test-retest reproducibility of the exercise treadmill examination in lumbar spinal stenosis. *Mayo Clin.Proc.* 2000;75:1002-7.

5. Yamashita K, Hayashi J, Ohzono K et al. Correlation of patient satisfaction with symptom severity and walking ability after surgical treatment for degenerative lumbar spinal stenosis. *Spine.* 2003;28:2477-81.

6. Barz T, Melloh M, Staub L et al. The diagnostic value of a treadmill test in predicting lumbar spinal stenosis. *Eur Spine J.* 2008;..

7. Tenhula J, Lenke LG, Bridwell KH et al. Prospective functional evaluation of the surgical treatment of neurogenic claudication in patients with lumbar spinal stenosis. *J.Spinal Disord.* 2000;13:276-82.

8. Atlas SJ, Keller RB, Wu YA et al. Long-term outcomes of surgical and nonsurgical management of lumbar spinal stenosis: 8 to 10 year results from the maine lumbar spine study. *Spine.* 2005;30:936-43.

9. Brussee P, Hauth J, Donk RD et al. Self-rated evaluation of outcome of the implantation of interspinous process distraction (X-Stop) for neurogenic claudication. *Eur Spine J.* 2008;17:200-3.

10. Deen HG, Zimmerman RS, Lyons MK et al. Use of the exercise treadmill to measure baseline functional status and surgical outcome in patients with severe lumbar spinal stenosis. *Spine.* 1998;23:244-8.

11. Deen HG, Jr., Zimmerman RS, Lyons MK et al. Measurement of exercise tolerance on the treadmill in patients with symptomatic lumbar spinal stenosis: a useful indicator of functional status and surgical outcome. *J Neurosurg.* 1995;83:27-30.

12. Fritz JM, Erhard RE, Delitto A et al. Preliminary results of the use of a two-stage treadmill test as a clinical diagnostic tool in the differential diagnosis of lumbar spinal stenosis. *J Spinal Disord.* 1997;10:410-6.

13. Adamova B, Vohanka S, Dusek L. Differential diagnostics in patients with mild lumbar spinal stenosis: the contributions and limits of various tests. *Eur Spine J.* 2003;12:190-6.

14. Whitehurst M, Brown LE, Eidelson SG et al. Functional mobility performance in an elderly population with lumbar spinal stenosis. *Arch Phys.Med Rehabil.* 2001;82:464-7.

15. Herno A, Airaksinen O, Saari T et al. Computed tomography findings 4 years after surgical management of lumbar spinal stenosis. No correlation with clinical outcome. *Spine.* 1999;24:2234-9.

16. Herno A, Partanen K, Talaslahti T et al. Long-term clinical and magnetic resonance imaging follow-up assessment of patients with lumbar spinal stenosis after laminectomy. *Spine.* 1999;24:1533-7.

17. Herno A, Saari T, Suomalainen O et al. The degree of decompressive relief and its relation to clinical outcome in patients undergoing surgery for lumbar spinal stenosis. *Spine.* 1999;24:1010-4.

18. Jensen OH, Schmidt-Olsen S. A new functional test in the diagnostic evaluation of neurogenic intermittent claudication. *Clin.Rheumatol.* 1989;8:363-7.

19. Moon ES, Kim HS, Park JO et al. Comparison of the predictive value of myelography, computed tomography and MRI on the treadmill test in lumbar spinal stenosis. *Yonsei Med J.* 2005;46:806-11.

20. Stucki G, Daltroy L, Liang MH et al. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine.* 1996;21:796-803.

21. Fairbank JC, Couper J, Davies JB et al. The Oswestry low back pain disability questionnaire. *Physiotherapy.* 1980;66:271-3.

22. Gepstein R, Arinzon Z, Adunsky A et al. Decompression surgery for lumbar spinal stenosis in the elderly: preoperative expectations and postoperative satisfaction. *Spinal Cord.* 2006;44:427-31.

23. Yamashita K, Ohzono K, Hiroshima K. Patient satisfaction as an outcome measure after surgical treatment for lumbar spinal stenosis: testing the validity and discriminative ability in terms of symptoms and functional status. *Spine.* 2006;31:2602-8.

24. Yamashita K, Ohzono K, Hiroshima K. Five-year outcomes of surgical treatment for degenerative lumbar spinal stenosis: a prospective observational study of symptom severity at standard intervals after surgery. *Spine.* 2006;31:1484-90.

25. Gelalis ID, Stafilas KS, Korompilias AV et al. Decompressive surgery for degenerative lumbar spinal stenosis: long-term results. *Int.Orthop.* 2006;30:59-63.

26. Shabat S, Arinzon Z, Folman Y et al. Long-term outcome of decompressive surgery for lumbar spinal stenosis in octogenarians. *Eur Spine J.* 2008;17:193-8.

27. Katz JN, Dalgas M, Stucki G et al. Degenerative lumbar spinal stenosis. Diagnostic value of the history and physical examination. *Arthritis Rheum.* 1995;38:1236-41.

28. Watters WC, III, Baisden J, Gilbert TJ et al. Degenerative lumbar spinal stenosis: an evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis. *Spine J.* 2008;8:305-10.

29. Swerts PM, Mostert R, Wouters EF. Comparison of corridor and treadmill walking in patients with severe chronic obstructive pulmonary disease. *Phys.Ther.* 1990;70:439-42.

30. Pearce ME, Cunningham DA, Donner AP et al. Energy cost of treadmill and floor walking at self-selected paces. *Eur J Appl Physiol Occup.Physiol.* 1983;52:115-9.

31. Katz JN, Lipson SJ, Chang LC et al. Seven- to 10-year outcome of decompressive surgery for degenerative lumbar spinal stenosis. *Spine.* 1996;21:92-8.

32. Katz JN, Lipson SJ, Lew RA et al. Lumbar laminectomy alone or with instrumented or noninstrumented arthrodesis in degenerative lumbar spinal stenosis. Patient selection, costs, and surgical outcomes. *Spine.* 1997;22:1123-31.

33. Iversen MD, Katz JN. Examination findings and self-reported walking capacity in patients with lumbar spinal stenosis. *Phys.Ther.* 2001;81:1296-306.

34. Katz JN, Stucki G, Lipson SJ et al. Predictors of surgical outcome in degenerative lumbar spinal stenosis. *Spine.* 1999;24:2229-33.

35. Frazier DD, Lipson SJ, Fossel AH et al. Associations between spinal deformity and outcomes after decompression for spinal stenosis. *Spine.* 1997;22:2025-9.

36. Simotas AC. Nonoperative treatment for lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;153-61.

37. Hansraj KK, O'Leary PF, Cammisa FP, Jr. et al. Decompression, fusion, and instrumentation surgery for complex lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;18-25.

38. Hansraj KK, Cammisa FP, Jr., O'Leary PF et al. Decompressive surgery for typical lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;10-7.

39. Lyle MA, Manes S, McGuinness M et al. Relationship of physical examination findings and self-reported symptom severity and physical function in patients with degenerative lumbar conditions. *Phys.Ther.* 2005;85:120-33.

40. Zucherman JF, Hsu KY, Hartjen CA et al. A multicenter, prospective, randomized trial evaluating the X STOP interspinous process decompression system for the treatment of neurogenic intermittent claudication: two-year follow-up results. *Spine.* 2005;30:1351-8.

41. Zucherman JF, Hsu KY, Hartjen CA et al. A prospective randomized multi-center study for the treatment of lumbar spinal stenosis with the X STOP interspinous implant: 1-year results. *Eur Spine J.* 2004;13:22-31.

42. Stucki G, Liang MH, Fossel AH et al. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J.Clin.Epidemiol.* 1995;48:1369-78.

43. Tomkins CC, Battie MC, Hu R. Construct validity of the physical function scale of the Swiss Spinal Stenosis Questionnaire for the measurement of walking capacity. *Spine.* 2007;32:1896-901.

44. Pratt RK, Fairbank JC, Virr A. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. *Spine.* 2002;27:84-91.

45. McGregor AH, Hughes SP. The evaluation of the surgical management of nerve root compression in patients with low back pain: Part 1: the assessment of outcome. *Spine.* 2002;27:1465-70.

46. Herno A, Airaksinen O, Saari T et al. Surgical results of lumbar spinal stenosis. A comparison of patients with or without previous back surgery. *Spine.* 1995;20:964-9.

47. Herno A, Airaksinen O, Saari T et al. Lumbar spinal stenosis: a matched-pair study of operated and non-operated patients. *Br.J.Neurosurg.* 1996;10:461-5.

48. Shabat S, Leitner Y, Nyska M et al. Surgical treatment of lumbar spinal stenosis in patients aged 65 years and older. *Arch.Gerontol.Geriatr.* 2002;35:143-52.

49. Herno A, Airaksinen O, Saari T. Computed tomography after laminectomy for lumbar spinal stenosis. Patients' pain patterns, walking capacity, and subjective disability had no correlation with computed tomography findings. *Spine.* 1994;19:1975-8.

50. Airaksinen O, Herno A, Turunen V et al. Surgical outcome of 438 patients treated surgically for lumbar spinal stenosis. *Spine.* 1997;22:2278-82.

51. Weinstein JN, Tosteson TD, Lurie JD et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. *N.Engl.J Med.* 2008;358:794-810.

52. Sirvanci M, Bhatia M, Ganiyusufoglu KA et al. Degenerative lumbar spinal stenosis: correlation with Oswestry Disability Index and MR Imaging. *Eur Spine J.* 2008;..

53. Haro H, Maekawa S, Hamada Y. Prospective analysis of clinical evaluation and self-assessment by patients after decompression surgery for degenerative lumbar canal stenosis. *Spine J.* 2008;8:380-4.

54. Zouboulis P, Karageorgos A, Dimakopoulos P et al. Functional outcome of surgical treatment for multilevel lumbar spinal stenosis. *Acta Orthop.* 2006;77:670-6.

55. Cavusoglu H, Turkmenoglu O, Kaya RA et al. Efficacy of unilateral laminectomy for bilateral decompression in lumbar spinal stenosis. *Turk.Neurosurg.* 2007;17:100-8.

56. Cavusoglu H, Kaya RA, Turkmenoglu ON et al. Midterm outcome after unilateral approach for bilateral decompression of lumbar spinal stenosis: 5-year prospective study. *Eur Spine J.* 2007;16:2133-42.

57. Walsh TL, Hanscom B, Lurie JD et al. Is a condition-specific instrument for patients with low back pain/leg symptoms really necessary? The responsiveness of the Oswestry Disability Index, MODEMS, and the SF-36. *Spine.* 2003;28:607-15.

58. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine.* 2000;25:2940-52.

59. Gronblad M, Hupli M, Wennerstrand P et al. Intercorrelation and test-retest reliability of the Pain Disability Index (PDI) and the Oswestry Disability Questionnaire (ODQ) and their correlation with pain intensity in low back pain patients. *Clin.J.Pain.* 1993;9:189-95.

60. Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain.* 1996;65:71-6.

61. Cronbach LJ, Meehl PE. Construct Validity in Psychological Tests. *Psychol.Bull.* 1955;52:281-302.

62. Feeny D, Furlong W, Boyle M et al. Multi-attribute health status classification systems. Health Utilities Index. *Pharmacoeconomics.* 1995;7:490-502.

63. Andresen EM. Criteria for assessing the tools of disability outcomes research. *Arch Phys.Med Rehabil.* 2000;81:S15-S20.

64. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin* 1959;56:81-105.

65. Lee J, Hida K, Seki T et al. An interspinous process distractor (X STOP) for lumbar spinal stenosis in elderly patients: preliminary experiences in 10 consecutive cases. *J.Spinal Disord.Tech.* 2004;17:72-7.

66. Sinikallio S, Aalto T, Airaksinen O et al. Somatic comorbidity and younger age are associated with life dissatisfaction among patients with lumbar spinal stenosis before surgical treatment. *Eur.Spine J.* 2006;..

67. Thornes E, Grotle M. Cross-cultural adaptation of the Norwegian version of the spinal stenosis measure. *Eur Spine J.* 2008;..

68. Johnsson KE, Rosen I, Uden A. The natural course of lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 1992;82-6.

69. Ogikubo O, Forsberg L, Hansson T. The relationship between the cross-sectional area of the cauda equina and the preoperative symptoms in central lumbar spinal stenosis. *Spine.* 2007;32:1423-8.

CHAPTER 7

GENERAL DISCUSSION AND CONCLUSIONS

7.1 Overview

Walking is a fundamental component of everyday life and is important for transportation, physical activities and health. To most people, mobility is perhaps the single activity of daily living on which they place the most value.[1] Measurement of walking capacity in LSS populations is vital for monitoring the natural progression of the condition and for examining outcomes of treatment. However, research regarding the validity of measures of walking capacity in LSS populations is limited. The overall purpose of this thesis research was to provide new validity evidence regarding the use of various measures of walking capacity in LSS. In the absence of an established criterion measure, it was proposed that the construct specific Self-Paced Walking Test (SPWT) is the best measure of walking capacity for this population. Yet it is likely that the SPWT may not be feasible in many research and clinical settings. Therefore, further research was warranted to examine the construct and criterion validity of the measures of walking capacity currently employed and reported in the LSS literature, including a treadmill protocol and various self-report instruments.

7.2 The Validity Argument

The process of forming a validity argument in this research involved explicitly defining the construct of interest (walking capacity in LSS), stating hypotheses regarding outcomes of, and expected relationships between, various measures of walking capacity, and using statistical tests to either confirm or dispute these theory-based hypotheses. A number of the major hypotheses in the present research were supported, providing a number of new and valuable pieces of validity evidence regarding the use of various measures of walking capacity in LSS. The major pieces of validity evidence include:

1. Construct validity evidence (both convergent and divergent) supporting the use of the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire for the measurement of walking capacity in persons with lumbar spinal stenosis.

2. Evidence confirming reproducibility of the chosen criterion measure of walking capacity (Self-Paced Walking Test), as well as the Physical Function Scale, the ODI and various single item self-report measures of walking capacity. Evidence regarding reproducibility of the SPWT confirmed the acceptability of the SPWT as a criterion.

3. Construct validity evidence regarding the use of a treadmill protocol and various single item self-report measures in the measurement of walking capacity in LSS through confirmed hypotheses regarding the expected relationships among measures of walking capacity (convergent), and measures of other unrelated constructs (divergent).

4. Criterion related validity evidence regarding the use of a treadmill protocol, the Physical Function Scale and various single item self-report measures as surrogate measures of walking capacity in LSS, using the Self-Paced Walking Test as the criterion measure. This evidence confirms that the construct being measured in the treadmill test and the self-report measures which were highly correlated with the criterion is likely walking capacity. However, despite the high correlation between the SPWT and treadmill test, evidence was provided to suggest that the treadmill test significantly under-estimates distance measured during the SPWT.

7.3 The SPWT as the Criterion Measure of Walking Capacity in LSS

As stated in Study 3 (Chapter 6), some readers may question the use of the SPWT as a criterion measure of walking capacity, given that it is a new method. However, as described previously, the fact that the SPWT is a direct operationalization of the construct of walking capacity in LSS provides evidence supporting its construct validity. The SPWT is relevant to and representative of the construct of walking capacity as defined in this thesis research. The SPWT was also found to be reproducible in Study 2 (Chapter 5), providing support for its utility as a criterion. The following sections outline the rationale surrounding the choice of the SPWT as the criterion measure of walking capacity for this research.

7.3.1 *Rationale for the Choice of Criterion Measure Used in the Present Research*

The construct of walking capacity has been defined as the distance a person with lumbar spinal stenosis is able to walk without support on a level surface at a self-selected speed before being forced to stop due to symptoms of LSS. The SPWT taps this construct as closely as possible. The test consisted of subjects walking around a track, without support, at their own pace, and until they were forced to stop due to symptoms of LSS (or other reasons), or until a time limit of 30 minutes had been reached.

7.3.2 *Choice of Speed for the Criterion*

In examining walking capacity, we were interested in how far subjects are able to walk in a realistic, everyday situations (e.g. if a person were to leave their home, how far would they be able to go before being forced to stop due to symptoms of LSS). Thus, it was decided that a self-selected speed was most appropriate given that imposing a set speed for all subjects may not accurately represent the speed at which a particular subject would walk in a real-life setting. It is likely that when walking for activities of daily living, persons with LSS vary their speed according to comfort level and symptoms. Thus, an un-regulated self-selected speed most closely represents the construct of walking capacity, as defined. Results from Study 2 (Chapter 5) of the present thesis confirm that self-selected speed during the SPWT is reproducible and stable over a period of four days up to three weeks.

7.3.3 *Choice of Track vs. Treadmill as the Criterion*

The use of a treadmill protocol as the criterion measure was considered. As mentioned, treadmill protocols have been used previously and found to be reliable for measuring walking in LSS populations.[2-5] However, we believe that walking on a track more closely represents walking in everyday, functional situations. It is likely that given the elderly nature and balance problems associated with LSS patients, a treadmill test may be uncomfortable and intimidating for many patients. In addition, it is more difficult for subjects to walk at a self-selected speed on a treadmill, given that they must use manual controls to adjust speed. Results of this thesis research confirm that a treadmill test can not be

used as a gold standard measure of walking in LSS, given that it systematically and significantly under-estimated patients' walking capacities, as well as walking speed and total time.

7.3.4 *Choice of Grade for the Criterion*

The choice of a level surface for testing was based on the pathology of spinal stenosis. It has been shown that spinal extension and neutral postures that are present in weight bearing activities such as walking may cause a relative narrowing of the spinal canal and exacerbate symptoms of LSS.[6] Conversely, spinal flexion may increase the dimensions of the spinal canal and decrease symptoms of LSS.[7;8] It has been shown that both an increase in walking grade and the use of handrails precipitate a forward flexed position, which may open the lumbar canal and relieve symptoms, thus rendering a test of LSS symptoms inaccurate.[2;6;9] Fritz et al. found that subjects walking at 15% grade showed an average increase in forward flexion of 8° and were able to walk significantly longer unlimited by symptoms compared to subjects walking on a level surface.[6] Thus, in order to precipitate symptoms of LSS in less time and provide and accurate test of LSS symptoms during unassisted walking, the choice of a level surface seemed appropriate.

7.3.5 *Test Duration of the Criterion*

All subjects for the present research had been diagnosed with LSS and associated walking limitations. However, unlike most other studies of walking in LSS, subjects in the present research were not all undergoing surgery for LSS. Thus, it was anticipated that symptoms would potentially be less severe and walking capacities greater than those of subjects who participated in previous LSS walking studies (using treadmill protocols). For this reason, we chose a time limit of 30 minutes, as opposed to the 15-20 minute limits used for treadmill protocols by Deen et al.[2;3], Herno et al.[10-12], and Tenhula et al.[13] Tenhula et al. reported a mean treadmill walking time of 15.3 minutes ±6 minutes in patients prior to surgery for LSS.[13] This implies that even patients with stenosis severe enough to warrant surgery are capable of walking 15 minutes or more. Thus, the increase in time limit to 30 minutes allowed more time for symptom precipitation in patients

with less severe symptoms, and was intended to prevent a ceiling effect similar to that observed by Deen et al. with a 15 minute limit.[2]

Results of the present research suggest however, that even 30 minutes may not be a sufficient time limit, given that 10 of 41 (24%) subjects were able to walk for the full 30 minutes (Study 3, Chapter 6). It is possible that a walking test of unlimited time may be necessary to evaluate less limited LSS patients. One may question whether or not it is important to measure walking capacity in patients who are able to walk for 30 minutes, as they may be considered 'unlimited'. However, all of the subjects who walked for 30 minutes *did* report symptoms of LSS which increased with walking, suggesting that if they were allowed to walk longer, they would eventually reach a symptom limited end point. The important question to consider here is what a patient deems to be limiting for them. A patient who could run 10km prior to the onset of stenosis symptoms would feel very limited if they are now only able to walk for 35 minutes. A walking test of unlimited time would therefore likely be very useful in evaluating the natural progression of the condition, or outcomes of treatment for patients whose symptoms are less severe and whose walking capacities are less limited. Outcomes of non-surgical therapies for stenosis could be evaluated using a walking test of unlimited time. However, one would have to weigh the benefits with the practicalities of conducting such a test. As mentioned in Chapter 6, it is also possible that other options for measuring walking capacity in less limited LSS patients could be considered, which would allow for patients to measure their own capacities in realistic day to day settings at home. Pedometers with distance measuring capabilities are a possibility for this type of measurement.[14]

7.3.6 *Reproducibility of the Criterion and Stability of Walking in LSS Patients*

In correlational studies of validity, reproducibility cannot be overlooked, given that the maximum possible correlation between two measures is the square root of the product of their reliabilities. Therefore, it was necessary to examine reproducibility of the criterion SPWT, so as to determine whether validity coefficients were subject to attenuation because of low reliability. It was also necessary to examine reproducibility to ensure that the SPWT would be able to

detect actual changes in walking capacity over time. Reproducibility concerns the degree to which repeated measurements in relatively stable study objects, yield similar results.[15] Given that LSS is a chronic condition, with rare acute symptom exacerbations, it is reasonable to expect that walking capacity would remain stable over a relatively short period of time. Indeed, results of the present research demonstrated that at the group level, distance and speed measured in the SPWT are highly reproducible over a period of four days up to three weeks, implying that walking capacity may also be stable over this period. These results are consistent with studies examining the natural history of LSS, which suggest that walking capacity in LSS remains largely unchanged over time, for up to 10 years.[17] Yet, it has been suggested that walking capacity may vary from day to day in LSS patients.[16] This notion was supported at the individual level, where the SPWT was found to be less reproducible. There were a few patients who demonstrated substantial changes in their walking capacity between tests, suggesting that multiple tests of walking capacity should be taken if accuracy of measurement is important.

As previously discussed, there were two subjects in Study 2 (Chapter 5) who walked substantially less during their second SPWT. However, both of these subjects did indicate that their walking capacity had deteriorated over the 14-15 days between tests. This implies that while on the whole walking capacity is stable over up to 3 weeks, there are patients who may experience changes in their condition over a period as short as two weeks. In addition there was one subject who walked more than 700% further during the second SPWT. This is likely because her first SPWT gave her the confidence to walk further without her walker during the second test. This suggests that patients who are dependent on walking equipment for mobility may need an orientation session before testing to ensure that no learning effect is associated with the use of the walking test.

To provide additional information regarding the stability of walking in LSS patients, we examined whether subjects could gauge changes in their walking capacities between administrations of the SPWT using a change score. Before the 2nd SPWT, subjects were asked whether walking capacity was about the same, a

little bit better/worse, moderately better/worse or a lot better/worse compared to the first SPWT. Interestingly there was no significant relationship observed between the walking capacity change score and differences in distance observed between SPWT administrations. This suggests that patients' perceptions of changes in their walking capacity are not accurately reflected by changes in walking distance during the SPWT. In addition, the walking capacity change score was not significantly correlated with changes in worst pre-walking test pain, or the ODI. This suggests that patient perceived changes in walking capacity are not related to changes in pre-walking pain or back-pain related disability. These results have clinical implications, given that many clinicians judge patient status based on questions regarding change. It is possible that measuring perceived changes in walking capacity using a self-report change score is not sufficient to capture actual changes in walking capacity, as measured using an observational test.

7.3.7 *Validity of the Criterion (SPWT)*

The SPWT was used as the criterion measure in the present research, given that it is thought to have construct validity as a direct operationalization of the construct of walking capacity in LSS. Results of Studies 2 and 3 (Chapters 5 and 6) provided evidence regarding the construct validity of the SPWT. Study 2 found the SPWT to be reproducible, while in Study 3, the hypothesized magnitude and pattern of correlations with measures of similar and different constructs was supported, providing both convergent and divergent validity evidence. No studies have been conducted prior to this research examining criterion validity of measures of walking in LSS, and no realistic alternative criterion measures of walking capacity have been suggested in the literature. Walking distance on a treadmill has been used as a gold standard measure of walking capacity in one study by Moon et al.,[5] yet for reasons previously mentioned the SPWT is preferable to treadmill testing of walking in LSS.

7.4 Treadmill Testing of Walking Capacity in LSS

7.4.1 *Rationale for Choice of Treadmill Protocol*

The treadmill protocol in the present research was a modification of the protocol described by Deen et al.[2-4] Deen's protocol is highly cited in LSS research and has the most information available regarding its design and reproducibility. This protocol was adjusted to be as closely matched to the SPWT as possible. Thus, changes to Deen's protocol included extending the duration from 15 minutes to 30 minutes and using only self-selected speeds. It was reasonable to use only a self-selected speed trial, given that there were no significant differences observed for distance or time between the 1.2mph and self-selected speed trials in any of Deen's studies.[2-4] Rationales for the selected test parameters are the same for the treadmill test as for the SPWT (self-selected speed, 0% grade and a 30 minute time limit). Similar to the protocol described by Deen et al.[2] we also chose to have subjects maintain an upright posture, and avoid using handrails to prevent flexion based alleviation of symptoms.

7.4.2 *Validity of the Treadmill Test*

This is the first study of which we are aware designed to examine validity of a treadmill test for the measurement of walking in LSS. Both construct and criterion validity evidence was provided to support the use of the treadmill test for this purpose. However, although a high correlation was observed between the SPWT final distance and the treadmill final distance, there was a systematic bias, with subjects walking significantly further in the SPWT. On average, the treadmill test under-estimated the SPWT distance by 25%-35%, with the majority of subjects (63%) walking at least 25% further during the SPWT. There are a number of factors that could explain this observation, including walking speed, comfort level, balance and walking time. Although patients were allowed to select and modify their speed while walking on the treadmill, it is likely that the chosen speeds are not representative of their normal level ground walking. This is evidenced by the fact that mean walking speed on the treadmill (2.5km/h ± 1.0 km/h or 1.5mph ± 0.6 mph) was significantly ($p<0.05$) slower than that on the SPWT (4.1km/h ± 1.1km/h or 2.5mph±0.7mph). It is likely that patients selected

a speed that is slower than their habitual speeds due to uncertainty regarding balance and comfort level on the treadmill. Treadmill walking is by no means natural. The surface is moving for you, meaning that you must keep up with the pace of the treadmill, as opposed to being able to make minor modifications to walking speed based on symptoms, as one could do on solid ground.

In addition to speed, total walking time was found to be significantly ($p<0.05$) less for the treadmill (771seconds ± 673 seconds) than for the SPWT (912 seconds ± 666 seconds). Excluding the ten subjects who walked for the full 30 minutes on the SPWT and treadmill tests, all tests were symptom limited. This implies that subjects may actually be experiencing severe symptoms sooner using a treadmill test, causing them to stop the test earlier than they would during a SPWT. It is also possible that subjects may be stopping the test in part due to reasons other than severe symptoms, such as being uncomfortable or feeling unsteady. Although all subjects that did not walk the full 30 minutes reported stopping due to symptoms of LSS, it is possible that there are other contributing factors which the subjects did not want to admit or report. It is likely that a combination of severe symptoms and unease or lack of balance on the treadmill lead to subjects to terminate the test earlier on the treadmill than on solid ground. It is possible that an attempt to make treadmill walking more comfortable could narrow the gap between treadmill and level ground walking. This could be accomplished by using treadmills which are level with the ground and thus more natural to walk on. In addition, it is possible that boredom played a role in patients' shorter walking time on the treadmill compared to the track. Perhaps inserting TV's or screens with advancing landscapes and music would aid in making the treadmill test more engaging.

The aforementioned factors provide more evidence to support the SPWT as a criterion measure of walking in this population. The SPWT provides a more realistic walking environment and setting, where subjects can walk at their own pace and adjust their speed as needed. The SPWT elicits a real symptom limited test, as opposed to a test potentially limited by unease or intimidation. Therefore, when using a treadmill test in research and clinical settings, test administrators

should be aware that the test is likely to substantially under-estimate patients walking capacities (distance) as well as walking speed and time.

7.5 Physical Function Scale of the Swiss Spinal Stenosis Questionnaire

Information regarding the use of the Physical Function Scale in LSS populations is of high priority given the recent NASS clinical guidelines for LSS which suggested that the Swiss Spinal Stenosis Questionnaire is currently the best and most specific outcome measure for use with LSS populations.[18] The same guidelines also suggested that in future studies focusing on specific outcome measures for the treatment of LSS, the Swiss Spinal Stenosis Questionnaire could potentially be considered a gold standard.[18] However, results of the present study do not support the use of the Physical Function Scale as a gold standard measure of walking capacity in LSS.

The range of mean Physical Function Scale scores over the three studies in the present thesis research was 2.1-2.3. This range of scores suggests slightly less severity on average than reported in most previous studies of LSS, which included only subjects who were undergoing surgery. In a study of LSS surgical outcomes, Katz et al. reported pre-operative Physical Function Scale scores ranging from 3.5-3.7 and post-operative scores ranging from 0.9-1.5.[19] Lee et al. reported a mean pre-operative score of 2.7 and a post-operative score of 2.2,[20] while Sinikallio et al.[21] reported a mean pre-operative score of 2.5. In two studies examining the X-STOP surgical procedure for LSS, Zucherman et al. reported a mean pre-operative score of 2.5±0.5 and a range of 1.6-3.6.[22;23] The only study of which we are aware which reported a mean Physical Function Scale score for non-surgical patients is that of Lyle et al. [24] who reported a score of 2.0±0.6 in a group of low back pain patients, 50% of whom had LSS.

7.5.1 *Reproducibility of the Physical Function Scale*

Given that all other studies examining reproducibility of the Physical Function Scale have been conducted with surgical patients, results of the present research provide new information, given that we used a more heterogeneous sample of LSS patients. Study 2 (Chapter 5) of the present research suggests that the Physical Function Scale has acceptable reproducibility, with an ICC of 0.77

(95% confidence interval, 0.56-0.89), a mean difference between test administrations of 0.04 and an MAD of 0.3. Although the value of the ICC is slightly lower than expected in the current research, results of acceptable Physical Function Scale reproducibility are substantiated by Pratt et al.[25] who reported an ICC of 0.82 in a sample of 29 LSS patients with neurogenic claudication, as well as Thornes et al. who reported an ICC of 0.89 (95% confidence interval, 0.79-0.95) in a sample of 75 patients prior to LSS surgery.[26] Stucki et al. also reported a Spearman correlation of 0.94 between test-retest administrations of the Physical Function Scale in 23 surgical patients. It appears from the literature that the Physical Function Scale is reproducible in patients undergoing surgery for LSS. Results of the present research suggest that it is also reproducible in those whose LSS conditions may not be severe enough to warrant surgery.

7.5.2 *Validity of the Physical Function Scale*

Prior to the present thesis research, no studies have been conducted examining the validity of the Physical Function Scale for the measurement of walking capacity in LSS. The only pieces of validity evidence available to support the use of the Physical Function scale in the assessment of walking capacity in LSS were from a study by Stucki et al.,[27] the scale developers, who reported correlations between the Physical Function Scale and physicians' assessment of walking capacity (r=0.47), as well as the physical dimension of the Sickness Impact Profile (SIP) (r=0.49), in a sample of LSS patients undergoing surgery.[28] The present thesis research has added much to the knowledge base surrounding the use of the Physical Function Scale for the measurement of walking capacity in LSS.

Study 1 (Chapter 4) provided both convergent and divergent construct validity evidence regarding the use of the Physical Function Scale for the measurement of walking in LSS.[29] Yet, it could not be ascertained from Study 1 that the construct being measured was in fact walking capacity. Using a different sample of LSS patients, Study 3 (Chapter 6) provided criterion related validity evidence for the Physical Function Scale, confirming that the construct being measured using this scale is likely walking capacity. Results of Study 3 also

confirm construct validity conclusions from Study 1, as both convergent and divergent validity evidence was provided. Correlations with the Physical Function Scale were similar between Studies 1 and 3 for the ODI (r=0.72 and r=0.65), ODI walking item (r=0.80 and r=0.80), HUI3 (r=0.61 and r=0.63) and the HUI3 ambulation score (r=0.62 and r=0.56), suggesting that results can be generalized across the two sample populations studied. In addition, the walking distance specific item from the Physical Function Scale was found to be acceptably reproducible, and more highly correlated with the criterion (r=0.72) than the scale as a whole (r=0.68), suggesting that this item may be used in isolation to examine walking distance specifically.

Although the Physical Function Scale has been used primarily to assess outcomes of surgical interventions for LSS patients,[19;22;23;30-35] results of the present study indicate that the scale is both reproducible and valid for use with LSS patients of varying severities, including those who do not elect to have surgery. Future studies of conservative treatment for LSS could benefit from using this condition specific tool in their assessment methods, if self-report measures are desired.

7.6 Oswestry Disability Index

7.6.1 *Reproducibility of the Oswestry Disability Index*

Prior to this research, reproducibility of the ODI has only been examined once in an LSS population. Pratt et al. reported an ICC of 0.89 in a group of patients with neurogenic claudication who were undergoing surgery for LSS.[25] Results of Study 2 (Chapter 5) are consistent with this, suggesting that the ODI is reproducible in LSS patients of varying severities (ICC= 0.90, mean difference between administrations of 0.8% and MAD of 4.1%). The single item addressing pain limited walking capacity from the ODI was also found to be reproducible in the present study, suggesting that it may be possible to use this item in isolation, if users of the ODI are interested in walking capacity specifically.

7.6.2 *Validity of the Oswestry Disability Index*

Prior to the present research, the validity of the ODI for the measurement of walking in LSS had received very limited investigation. Given the widespread

use of the ODI in LSS and low back pain research, it is valuable to determine if the ODI is valid for use in evaluating walking capacity, to allow for comparison of walking capacity between LSS studies and between different low back conditions.

The range of mean ODI scores from the three studies in the present thesis research (43-47%) is within the range of scores reported in the literature for LSS populations.[36-38] The only validity evidence to date regarding the use of the ODI in measuring walking in LSS includes a correlation of (r= 0.51) between the ODI and total distance walked during a treadmill test,[39] and a significant correlation between post-operative ODI scores and treadmill test parameters (correlation coefficients not reported).[13] The results of Study 3 (Chapter 6) are consistent with the literature, and suggest that the ODI as a whole is not an appropriate instrument for the measurement of walking capacity in LSS, given the relatively low correlations observed with both the criterion SPWT (r=0.41) and the treadmill test (r=0.40). However, given that many authors will continue to use the ODI because of its established reproducibility[25] and utility in multiple low back disorder populations, there may be another option for examining walking capacity using this instrument. One question in the ODI is specific to walking distance, and was found to be reproducible (ICC=0.89) and highly correlated with the SPWT (r=0.82), the treadmill test (r=0.83) and the Physical Function Scale (r=0.80). It is suggested that if the ODI is the instrument of choice in a study of LSS, the walking specific question alone may be a valid and reproducible means to evaluate walking capacity.

7.7 Self-reported Walking Distance Reproducibility and Validity

In order to determine whether people with LSS could accurately estimate their walking capacities, a construct specific question was designed for this research: "If you were to go for a walk today, how far would you be able to walk at your own pace, on level ground before being forced to stop due to symptoms of lumbar spinal stenosis (m)?" Subjects were asked this question with a visible distance reference of a 60m straight away on a track. Both construct and criterion validity evidence was provided for the use of this question in measuring walking

capacity, as it was convergently correlated with the SPWT (r=0.65), the treadmill (r=0.73) and the Physical Function Scale (r=0.57). However, we found that patients who did not walk for the full 30 minutes tended to underestimate their walking capacity as measured using the criterion SPWT by approximately 30%. On comparison with other self-report measures, correlations observed between this question and the other measures of walking are not as high as those observed for the Physical Function Scale, the Physical Function Scale distance specific item, or the ODI walking distance item. In addition, of the walking measures examined, this question showed the worst reproducibility, with an ICC of 0.50 (95% confidence interval, 0.17-0.74). Therefore, although this construct specific question may be valid for measuring walking distance in LSS patients, it is no more valid or reproducible than the already established measures used with LSS patients, including the Physical Function Scale. The only advantage to using this question in place of an ordinal scale (Physical Function Scale) is that the construct specific question is measured on a continuous scale, allowing for use of traditionally more powerful parametric statistics.

7.8 Walking Difficulty Scale Reproducibility and Validity

The 100mm visual analog scale (VAS) used in the present research came from work by Yamashita et al.[40] who examined the relationship between post-surgical patient satisfaction and self-reported walking difficulty. In Yamashita's study, self-reported difficulty walking was found to be the strongest, and only independent predictor of patient satisfaction (Spearman r=0.43) when the variables of postoperative back pain, leg pain, numbness and walking time were also considered.[40] However, results of the present study demonstrated that this item was not highly correlated with the criterion SPWT, or with the other measures of walking capacity in LSS. In addition, this item was not found to be reproducible. As such, it is suggested that this item not be used to examine walking capacity in LSS populations.

7.9 Characteristics of the Study Samples

Most of the subjects included in prior studies investigating measures of walking in LSS were undergoing surgery.[4;27;28] Conversely, we chose to utilize

more heterogeneous samples of patients seeking care for LSS and to include patients with varying degrees of severity. In this way we aimed to provide validity and reproducibility evidence regarding measurement of walking capacity in patients with LSS of varying severities, including those who are not necessarily limited to the degree that they elect surgery.

Subjects for this thesis research included two different samples of LSS patients (Table 7-1). One sample (n=72) was used for Study 1 (Chapter 4), and the other sample (n=41) for Studies 2 and 3 (Chapters 5 and 6) (Study 2 included 28 subjects from the sample of 41). All participants in this thesis research were seeking care for LSS. Subjects were all greater than 45 years of age and had LSS confirmed on imaging and by a spine specialist surgeon. Subjects in Studies 2 and 3 were required to have self-reported LSS associated walking limitations, while those in Study 1 were not. All subjects may or may not have been having surgery for LSS. Subjects for Study 1 were part of a multi-centre prospective longitudinal study of prognostic factors and outcomes of LSS, while the other sample (Studies 2 and 3) were referred by spine specialist surgeons for this research.

The heterogeneity of these samples is evidenced by examining the wide range of scores for the Physical Function Scale and ODI in all three studies, as well as the range of walking distances measured in the SPWT for Studies 2 and 3 (Chapters 5 and 6) (Table 7-1). The Physical Function Scale mean scores were quite similar over the three studies, ranging from 2.1-2.3. The range of individual scores reported for the Physical Function Scale was similar across all three studies, with values reaching from 1 (mild limitation) to 3.6 (severe limitation). Similarly, the ODI mean scores were very similar across studies, with values ranging from 43% to 48%. Individual ODI scores ranged from 19% (little back related disability) to 81% (severe back related disability) in all three studies. The walking distance measured in the SPWT during Studies 2 and 3 also spans a large range (11-3077m) indicating the presence of a wide variety of patient limitations and severities in the sample. Although we did not observationally measure walking in Study 1 (Chapter 4), the walking specific item from the ODI indicates that patients self-reported walking distances range from the minimum of 1 (in bed

most of the time) to 5 (pain does not prevent me walking any distance). The wide range of scores for these various measures indicates that the results of this research may be generalized across the spectrum of patients with varying symptom severities and walking limitations associated with LSS.

In order to ensure that results were consistent across the range of LSS patients and limitations, additional post-hoc validity and reproducibility analyses were conducted. For Study 1 (Chapter 4), the inclusion of subjects with no reported walking limitations may have had the potential to affect the study results and generalizability. Yet when we conducted an additional analysis comparing the correlation matrix of subgroups of subjects who had reported walking limitations due to LSS (n=45) and those that did not (n=27), results were similar. This suggests that results are supportive of construct validity of the Physical Functions Scale over a range of self-reported walking limitations. For the reproducibility and validity analyses in Studies 2 and 3 (Chapters 5 and 6), subjects who walked for the full 30 minutes were removed and all analyses were repeated. Similar results were obtained for both groups, in both studies. This indicates that results can be generalized across the wide range of patients included in this research. The strength of the conclusions from these sub-group analyses is limited however by small sample sizes. More, these analyses were conducted to look for any systematic variations in results according to severity of walking limitation. Further research is warranted to repeat these studies with larger sample sizes, including larger sub-groups of LSS patients, based on degree of walking limitation.

7.10 Limitations

The primary limitation of Studies 2 and 3 (Chapters 5 and 6) was subject recruitment. The initial sample size estimate was 60, based on potential use of factor analysis with 5-10 subjects per variable (p.603).[41] Given an expected attrition rate of 20%, we were aiming to recruit 72 subjects. It was initially estimated by the primary spine surgeon collaborating on this project that he would personally see 8-10 patients with LSS per month, suggesting that this sample size

would be reasonable to obtain in one year. However, subject recruitment continued for over 2 years, with a sample size of 41 to date.

Given that factor analysis was not employed in this research, sample size calculations were adjusted according to recommendations with respect to correlation coefficients. It is suggested that appropriate sample sizes for correlational studies be determined based on confidence intervals around expected r values, with the goal of the confidence interval width being <0.2. With an expected correlation of 0.6 for convergent validity evidence, we would have required approximately 100 subjects to bring the confidence interval width below 0.2. However, the smaller than anticipated sample size did not appear to have any noticeable impact, given that the magnitude and significance of the validity coefficients and ICCs were as expected. According to Cohen,[42] with more than 40 subjects and correlation coefficients >0.60, as in both Studies 1 and 3 (Chapters 4 and 6), estimated power is greater than 0.99. With 28 subjects, as in Study 2 (Chapter 5) the power is 0.97 when correlations are >0.60.[42]

In addition to the smaller than anticipated sample size, the nature of the sample was limited by patient volunteering. Of 91 potential subjects who were referred by the spine specialist surgeons, 41 (45%) volunteered to participate for two visits, with only 28 (31%) returning for a third visit. I cannot comment on the characteristics of the patients who did not volunteer, as no information was collected from them. Given that subjects were required to come to the University of Alberta to participate, many declined to volunteer due to lack of transportation. Due to the large geographic region under Capital Health, many potential subjects lived hours away and as such were unable to come to Edmonton for repeat visits. There were also potential subjects living in the Edmonton area that did not have vehicles, or friends/family who could drive them to the University. While a number of subjects did arrive by public transport others were unwilling to make the trip using transit. Thus the sample was limited to subjects who were willing to come to the university and close enough geographically to have feasible transportation options.

Samples in all studies were limited by the source of patients. All patients in Studies 2 and 3 (Chapters 5 and 6) were referred by spine specialist surgeons in the Edmonton area because they were seeking consultation from a spine specialist for surgical care. It is possible that results would vary if patients were recruited from physiotherapists or other health professional practices where patients may not have been limited to the degree that they were seeking potential surgical intervention. Similarly, the sample in Study 1 (Chapter 4) was limited to patients who were seeking care for suspected stenosis in the Calgary region. Although all subjects had LSS confirmed on imaging, they were not all necessarily seeing spine specialist surgeons. This fact could explain why both the Physical Function Scale and ODI scores were slightly lower in the Study 1 sample compared to the sample for Studies 2 and 3.

7.11 Future research

7.11.1 *Reproducibility of Treadmill Testing in LSS Populations*

Given that we only collected data for one administration of the treadmill test, we were unable to examine its reproducibility. However, a number of investigators have reported on the reproducibility of similar treadmill tests.[4;5] Deen et al. reported concordance correlation coefficients for test re-test administrations of a treadmill testing in trials at patients' preferred speeds (CCC=0.96) and 1.2mph (CCC=0.89).[4] Moon et al. reported a Spearman correlation of 0.92 for test re-test reliability of the same Deen protocol, at 1.2mph.[5] However, the use of correlation coefficients to assess reproducibility has been questioned, given that correlation measures the strength of a relation between two variables in terms of agreement among ranked positions and not the agreement between the actual scores (distances, time).[43] Therefore, if the treadmill test investigated in the present research is to be used clinically, reproducibility should be examined using appropriate tests of reproducibility and agreement, such as an ICC or mean absolute deviation (MAD). Further research is also warranted to examine the reproducibility of treadmill testing in LSS to ensure that changes in walking capacity can be detected using this method.

7.11.2 *Validity and Reproducibility of Measures of Walking in LSS Subgroups*

Results of the present research identified subgroups of LSS patients, based on walking capacity (distance). Further research is warranted to investigate both validity and reproducibility of measures of walking in subgroups of patients with different degrees of walking limitation. It is possible that certain measures of walking are more valid and reproducible in specific subgroups of LSS patients. However, the results of the present research do suggest that the treadmill protocol and various self-report instruments were both reproducible and valid for use with LSS patients of varying degrees of walking limitations.

7.11.3 *Predictors of Walking Capacity in LSS*

It has been established that walking capacity is an important variable to monitor for assessment of treatment outcomes in LSS. It would also be very valuable to determine what factors predict walking capacity in LSS patients (e.g. pain, level of stenosis, distance to first symptoms, walking speed) in order to aid in setting priorities for treatment. One focus of such research would be the predictive value of anatomical stenosis severity identified with imaging. As previously discussed, the relationship between clinical outcomes (including walking capacity) and stenosis identified on imaging is unclear. Much of the literature in this area suggests that there is little or no relationship between anatomical imaging findings and clinical symptoms of LSS,[5;10;12;14;44-58] while some studies suggest that such a relationship does exist.[30;59-66] An investigation is planned to investigate predictors of walking capacity in LSS, including degree of anatomical stenosis measured on MRI, once data for 60 subjects has been obtained (10 subjects per 6 variables anticipated in the regression model).

In addition to examining predictors of walking capacity itself, future research could also examine predictors of changes in walking capacity over a short period of time. Study 2 (Chapter 5) demonstrated that some subjects' walking capacities may have actually changed substantially between re-tests of the SPWT. It would be valuable to determine what is actually causing the change in walking capacity over a period of up to 3 weeks. It is likely that these changes in walking capacity are due to changes in symptom severity. It is also possible

that these changes are due to factors such as motivation and boredom. Potentially the addition of a third SPWT in examining reproducibility would provide insight into whether or not changes in walking capacity are due to an actual improvement or deterioration in walking capacity, or simply due to natural day to day fluctuations.

7.11.4 *Responsiveness of Measures of Walking in LSS*

It would be valuable to examine the responsiveness of these walking capacity measures for LSS in order to determine which instruments or measures are best able to detect important clinical changes in walking capacity.[27] If walking capacity measures are chosen as primary outcomes in assessment of LSS treatments, knowledge of responsiveness would aid in selection of measures.[67] In addition, the determination of a minimal clinically important difference for walking capacity may be valuable in examining treatment outcomes. However, minimal clinically important difference may vary depending on patients' pre-treatment walking capacities and desire for change in walking capacity.

7.11.5 *Community-Based Measurement of Walking Capacity*

The criterion measure designed for this thesis research was based on the construct definition of walking capacity as the distance a person with LSS is able to walk without support on a level surface at a self-selected speed before being forced to stop due to symptoms of LSS. This criterion was selected in order to allow for standardized and reproducible measurement of walking in a feasible clinical setting, while attempting replicate real life walking conditions and allowing comparison with treadmill testing. However, it is possible that other clinicians or researchers may be interested in a slightly different construct, such as community-based walking capacity. It may be valuable to observe how far patients are able to walk in their natural, authentic home environment or neighbourhood. Measurement of walking in this way may be especially valuable in patients who are less limited. However, it is not practical for most researchers or clinicians to personally evaluate walking in this way due to time and personnel constraints, weather, walking surface concerns and lack of feasible measuring tools. It may be possible to identify community-based measures of walking

capacity in LSS patients, which can be administered by the patients themselves. Such tests would need to be easy to administer, while capturing valid walking capacity values. One potential option is the use of pedometers with distance measuring capabilities. Pedometers could also be used to monitor habitual daily activity, and volume of walking over longer periods of time (one day, one week). In fact, in the first study of its kind, Geisser and colleagues recently reported the use of pedometers to measure total distance walked over one week in patients with LSS.[14] Monitoring the distance a patient walks over one day, or one week may provide relevant information to clinicians regarding day to day patient walking needs and changes in walking. However, there would be concerns with reproducibility of such tests, given that they would be conducted in non-standardized and potentially changing environments, by patients themselves, as opposed to health professionals.

7.12 Conclusions

Overall, the results of the present thesis research provide validity evidence regarding the use of a number of measures of walking capacity in LSS, including a treadmill protocol, the Physical Function Scale of the Swiss Spinal Stenosis Questionnaire and the Oswestry Disability Index. These results have implications for both clinical and research environments. Validity evidence was provided supporting the use of a treadmill test for the measurement of walking in this population. However caution in using the treadmill test is warranted, given that it appears to significantly under-estimate patients' actual walking capacities by approximately 25%-35%, on average. In terms of self-report measures, validity and reproducibility evidence was provided supporting use of the Physical Function Scale for the measurement walking capacity. Given this evidence, along with condition specific nature and widespread use of this scale in LSS literature, the Physical Function Scale is purported to be an appropriate self-report instrument for use in the measurement of walking capacity in LSS. The walking distance specific item from the Physical Function Scale could be isolated to examine walking capacity with an even greater degree of accuracy than the scale in its entirety. The ODI as a whole was not found to be a good measure of

walking capacity in the present study. However, given the relationship observed between the walking specific item from the ODI and the criterion SPWT (as well as other measures of walking capacity) it is likely that this item could be isolated to investigate walking capacity in LSS populations when the ODI is the outcome instrument of preference. Validity evidence was provided supporting the use of the walking capacity construct specific question. However patients tended to under-estimate their actual walking capacities, and reproducibility of this item was not found to be acceptable. Overall, the results of this research provide valuable information for researchers and clinicians to aid in their selection of appropriate measures of walking capacity for use with LSS patients. Most importantly, this information is a valuable resource for individuals interpreting and using the results of the instruments and tests used in this thesis to measure walking capacity in people with LSS.

Table 7-1. Characteristics of the Study Samples

Variable	Study 1	Study 2	Study 3
Sample size	72	28*	41
Age (yrs)	69.5±11	66.8±8	66.9±10
Gender			
➢ Female (%)	51	57	57
Back pain present (%)	92	100	100
Duration of back pain (yrs)	9±11	11±12	10±11
Leg symptoms present (%)	85	100	100
Duration of leg pain (yrs)	7.5±9	8.5±11	7±9
Both back and leg symptoms present (%)	78	100	100
Walking limited due to back problem (%)	63	100	100
Outcome Measures			
Physical Function Scale (1-4)	2.1±0.7 (1-3.6)†	2.2±0.6 (1-3.4)	2.3±0.7 (1-3.4)
ODI (%)	43±14 (19-81)	48±13 (21-74)	47±13 (21-68)
SPWT distance (m)	NA	1142±967 (32-3077)	994 ±945 (11-3077)

* Subset of Study 3 sample
† Brackets indicate the range of scores

7.13 References

1. Chiou TJ, Tung SL, Wang WS et al. Pulmonary function changes in long-term survivors of chronic myelogenous leukemia after allogeneic bone marrow transplantation: a Taiwan experience. *Cancer Invest.* 2002;20:880-8.

2. Deen HG, Zimmerman RS, Lyons MK et al. Use of the exercise treadmill to measure baseline functional status and surgical outcome in patients with severe lumbar spinal stenosis. *Spine.* 1998;23:244-8.

3. Deen HG, Jr., Zimmerman RS, Lyons MK et al. Measurement of exercise tolerance on the treadmill in patients with symptomatic lumbar spinal stenosis: a useful indicator of functional status and surgical outcome. *J Neurosurg.* 1995;83:27-30.

4. Deen HG, Jr., Zimmerman RS, Lyons MK et al. Test-retest reproducibility of the exercise treadmill examination in lumbar spinal stenosis. *Mayo Clin.Proc.* 2000;75:1002-7.

5. Moon ES, Kim HS, Park JO et al. Comparison of the predictive value of myelography, computed tomography and MRI on the treadmill test in lumbar spinal stenosis. *Yonsei Med J.* 2005;46:806-11.

6. Fritz JM, Erhard RE, Delitto A et al. Preliminary results of the use of a two-stage treadmill test as a clinical diagnostic tool in the differential diagnosis of lumbar spinal stenosis. *J Spinal Disord.* 1997;10:410-6.

7. Schonstrom N, Lindahl S, Willen J et al. Dynamic changes in the dimensions of the lumbar spinal canal: an experimental study in vitro. *J Orthop.Res.* 1989;7:115-21.

8. Willen J, Danielson B, Gaulitz A et al. Dynamic effects on the lumbar spinal canal: axially loaded CT-myelography and MRI in patients with sciatica and/or neurogenic claudication. *Spine.* 1997;22:2968-76.

9. Fritz JM, Erhard RE, Vignovic M. A nonsurgical treatment approach for patients with lumbar spinal stenosis. *Phys.Ther.* 1997;77:962-73.

10. Herno A, Airaksinen O, Saari T et al. Computed tomography findings 4 years after surgical management of lumbar spinal stenosis. No correlation with clinical outcome. *Spine.* 1999;24:2234-9.

11. Herno A, Partanen K, Talaslahti T et al. Long-term clinical and magnetic resonance imaging follow-up assessment of patients with lumbar spinal stenosis after laminectomy. *Spine.* 1999;24:1533-7.

12. Herno A, Saari T, Suomalainen O et al. The degree of decompressive relief and its relation to clinical outcome in patients undergoing surgery for lumbar spinal stenosis. *Spine.* 1999;24:1010-4.

13. Tenhula J, Lenke LG, Bridwell KH et al. Prospective functional evaluation of the surgical treatment of neurogenic claudication in patients with lumbar spinal stenosis. *J.Spinal Disord.* 2000;13:276-82.

14. Geisser ME, Haig AJ, Tong HC et al. Spinal canal size and clinical symptoms among persons diagnosed with lumbar spinal stenosis. *Clin.J Pain.* 2007;23:780-5.

15. de Vet HC, Terwee CB, Knol DL et al. When to use agreement versus reliability measures. *J Clin.Epidemiol.* 2006;59:1033-9.

16. Porter RW. Spinal stenosis and neurogenic claudication. *Spine.* 1996;21:2046-52.

17. Amundsen T, Weber H, Nordal HJ et al. Lumbar spinal stenosis: conservative or surgical management?: A prospective 10-year study. *Spine.* 2000;25:1424-35.

18. Watters WC, III, Baisden J, Gilbert TJ et al. Degenerative lumbar spinal stenosis: an evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis. *Spine J.* 2008;8:305-10.

19. Katz JN, Lipson SJ, Lew RA et al. Lumbar laminectomy alone or with instrumented or noninstrumented arthrodesis in degenerative lumbar spinal stenosis. Patient selection, costs, and surgical outcomes. *Spine.* 1997;22:1123-31.

20. Lee J, Hida K, Seki T et al. An interspinous process distractor (X STOP) for lumbar spinal stenosis in elderly patients: preliminary experiences in 10 consecutive cases. *J.Spinal Disord.Tech.* 2004;17:72-7.

21. Sinikallio S, Aalto T, Airaksinen O et al. Somatic comorbidity and younger age are associated with life dissatisfaction among patients with lumbar spinal stenosis before surgical treatment. *Eur.Spine J.* 2006;..

22. Zucherman JF, Hsu KY, Hartjen CA et al. A multicenter, prospective, randomized trial evaluating the X STOP interspinous process decompression system for the treatment of neurogenic intermittent claudication: two-year follow-up results. *Spine.* 2005;30:1351-8.

23. Zucherman JF, Hsu KY, Hartjen CA et al. A prospective randomized multi-center study for the treatment of lumbar spinal stenosis with the X STOP interspinous implant: 1-year results. *Eur Spine J.* 2004;13:22-31.

24. Lyle MA, Manes S, McGuinness M et al. Relationship of physical examination findings and self-reported symptom severity and physical function in patients with degenerative lumbar conditions. *Phys.Ther.* 2005;85:120-33.

25. Pratt RK, Fairbank JC, Virr A. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. *Spine.* 2002;27:84-91.

26. Thornes E, Grotle M. Cross-cultural adaptation of the Norwegian version of the spinal stenosis measure. *Eur Spine J.* 2008;..

27. Stucki G, Liang MH, Fossel AH et al. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J.Clin.Epidemiol.* 1995;48:1369-78.

28. Stucki G, Daltroy L, Liang MH et al. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine.* 1996;21:796-803.

29. Tomkins CC, Battie MC, Hu R. Construct validity of the physical function scale of the Swiss Spinal Stenosis Questionnaire for the measurement of walking capacity. *Spine.* 2007;32:1896-901.

30. Katz JN, Lipson SJ, Chang LC et al. Seven- to 10-year outcome of decompressive surgery for degenerative lumbar spinal stenosis. *Spine.* 1996;21:92-8.

31. Iversen MD, Katz JN. Examination findings and self-reported walking capacity in patients with lumbar spinal stenosis. *Phys.Ther.* 2001;81:1296-306.

32. Katz JN, Stucki G, Lipson SJ et al. Predictors of surgical outcome in degenerative lumbar spinal stenosis. *Spine.* 1999;24:2229-33.

33. Frazier DD, Lipson SJ, Fossel AH et al. Associations between spinal deformity and outcomes after decompression for spinal stenosis. *Spine.* 1997;22:2025-9.

34. Simotas AC. Nonoperative treatment for lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;153-61.

35. Hansraj KK, Cammisa FP, Jr., O'Leary PF et al. Decompressive surgery for typical lumbar spinal stenosis. *Clin.Orthop.Relat Res.* 2001;10-7.

36. Airaksinen O, Herno A, Turunen V et al. Surgical outcome of 438 patients treated surgically for lumbar spinal stenosis. *Spine.* 1997;22:2278-82.

37. Herno A, Airaksinen O, Saari T et al. Lumbar spinal stenosis: a matched-pair study of operated and non-operated patients. *Br.J.Neurosurg.* 1996;10:461-5.

38. Herno A, Airaksinen O, Saari T et al. Surgical results of lumbar spinal stenosis. A comparison of patients with or without previous back surgery. *Spine.* 1995;20:964-9.

39. Barz T, Melloh M, Staub L et al. The diagnostic value of a treadmill test in predicting lumbar spinal stenosis. *Eur Spine J.* 2008;..

40. Yamashita K, Hayashi J, Ohzono K et al. Correlation of patient satisfaction with symptom severity and walking ability after surgical treatment for degenerative lumbar spinal stenosis. *Spine.* 2003;28:2477-81.

41. Tabachnick B, Fidell L. *Using Multivariate Statistics.* New York: Harper Collins, 1989:603.

42. Cohen J. *Statistical Power Analysis for the Behavioural Sciences.* 2nd ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1988.

43. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1:307-10.

44. Jalovaara P, Lahde S, Iikko E et al. The significance of residual stenosis after decompression for lumbar spinal stenosis. *Ann.Chir Gynaecol.* 1989;78:304-8.

45. Herno A, Airaksinen O, Saari T. Computed tomography after laminectomy for lumbar spinal stenosis. Patients' pain patterns, walking capacity, and subjective disability had no correlation with computed tomography findings. *Spine.* 1994;19:1975-8.

46. Postacchini F. Management of lumbar spinal stenosis. *J Bone Joint Surg.Br.* 1996;78:154-64.

47. Bernard TN, Jr. Repeat lumbar spine surgery. Factors influencing outcome. *Spine.* 1993;18:2196-200.

48. Amundsen T, Weber H, Lilleas F et al. Lumbar spinal stenosis. Clinical and radiologic features. *Spine.* 1995;20:1178-86.

49. Ciol MA, Deyo RA, Howell E et al. An assessment of surgery for spinal stenosis: time trends, geographic variations, complications, and reoperations. *J Am Geriatr.Soc.* 1996;44:285-90.

50. Frazier DD, Lipson SJ, Fossel AH et al. Associations between spinal deformity and outcomes after decompression for spinal stenosis. *Spine.* 1997;22:2025-9.

51. Hu RW, Jaglal S, Axcell T et al. A population-based study of reoperations after back surgery. *Spine.* 1997;22:2265-70.

52. Herron LD, Mangelsdorf C. Lumbar spinal stenosis: results of surgical treatment. *J Spinal Disord.* 1991;4:26-33.

53. Lehmann TR, Spratt KF, Tozzi JE et al. Long-term follow-up of lower lumbar fusion patients. *Spine.* 1987;12:97-104.

54. Resnick DK, Choudhri TF, Dailey AT et al. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 5: correlation between radiographic and functional outcome. *J Neurosurg.Spine.* 2005;2:658-61.

55. Lohman CM, Tallroth K, Kettunen JA et al. Comparison of radiologic signs and clinical symptoms of spinal stenosis. *Spine.* 2006;31:1834-40.

56. Uden A, Johnsson KE, Jonsson K et al. Myelography in the elderly and the diagnosis of spinal stenosis. *Spine.* 1985;10:171-4.

57. Weir B, de LR. Lumbar stenosis: analysis of factors affecting outcome in 81 surgical cases. *Can.J Neurol.Sci.* 1981;8:295-8.

58. Jonsson B, Annertz M, Sjoberg C et al. A prospective and consecutive study of surgically treated lumbar spinal stenosis. Part I: Clinical features related to radiographic findings. *Spine.* 1997;22:2932-7.

59. Ogikubo O, Forsberg L, Hansson T. The relationship between the cross-sectional area of the cauda equina and the preoperative symptoms in central lumbar spinal stenosis. *Spine.* 2007;32:1423-8.

60. Tuite GF, Doran SE, Stern JD et al. Outcome after laminectomy for lumbar spinal stenosis. Part II: Radiographic changes and clinical correlations. *J Neurosurg.* 1994;81:707-15.

61. McAfee PC, Ullrich CG, Yuan HA et al. Computed tomography in degenerative spinal stenosis. *Clin.Orthop.Relat Res.* 1981;221-34.

62. Porter RW, Ward D. Cauda equina dysfunction. The significance of two-level pathology. *Spine.* 1992;17:9-15.

63. Hamanishi C, Matukura N, Fujita M et al. Cross-sectional area of the stenotic lumbar dural tube measured from the transverse views of magnetic resonance imaging. *J Spinal Disord.* 1994;7:388-93.

64. Verbiest H. Results of surgical treatment of idiopathic developmental stenosis of the lumbar vertebral canal. A review of twenty-seven years' experience. *J Bone Joint Surg.Br.* 1977;59:181-8.

65. Beattie PF, Meyers SP, Stratford P et al. Associations between patient report of symptoms and anatomic impairment visible on lumbar magnetic resonance imaging. *Spine.* 2000;25:819-28.

66. Goh KJ, Khalifa W, Anslow P et al. The clinical syndrome associated with lumbar spinal stenosis. *Eur Neurol.* 2004;52:242-9.

67. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Control Clin.Trials.* 1991;12:142S-58S.

APPENDIX A
FORMS

INFORMATION SHEET

ASSESSMENT OF WALKING IN LUMBAR SPINAL STENOSIS
Principal Investigator(s): Christy Tomkins (Contact # 780.910.9318)
Supervisor: Dr. Michele Crites Battié

Co-Investigator(s): Dr. Harry Jiang, Dr. Stewart Petersen, Dr. Todd Rogers

Background: Due to a number of symptoms, people with lumbar spinal stenosis are often limited in their ability to walk. There are a number of tests and surveys available for looking at walking in people with lumbar spinal stenosis. But, we do not know how well these measures are assessing walking. Measurement of walking ability in people with spinal stenosis is important. Good measures of walking could be used by doctors to look at changes in the condition, before, and after treatment. Doctors might also assess walking in order to decide on treatment options.

Purpose: The purpose of this study is to look at how well the tests and surveys available to measure walking in people with lumbar spinal stenosis are actually measuring walking ability.

What will you have to do?: If you agree to take part, your participation in this study will involve two to three visits of about 1 hour each to the Van Vliet centre at the University of Alberta. There will be about 7 days between the visits. Total time for the study will be about 2.5 to 4 hours. During the first visit you will be asked to fill out a series of short surveys about your back problem and your overall health. Over the course of the three visits you will be asked to walk twice on the indoor track, and once on a treadmill. Both of these tests will be at your own pace for a maximum of 30 minutes.

Visit one (1.5 hours)
- Surveys, heart rate
- Walk either on the indoor track OR on the treadmill for up to 30 minutes, at your own pace.

Visit two (1.0 hour)
- Heart rate
- Walk either on the indoor track OR on the treadmill for up to 30 minutes, at your own pace.

Visit three (1.0 hour)
- Heart rate
- Walk on the indoor track for up to 30 minutes, at your own pace.

Total Time: About 4 hours

Will it help? Participation in this study will not have any direct effect on your condition. At the end of the study we will give you with the results of your surveys and walking test. You may want to share these results with your doctor. Your doctor may use these results to better evaluate your condition and treatment options. Your doctor may also use the results as a baseline to judge future changes. Information gained from this study will help surgeons and other health professionals in better caring for people with back problems such as yours.

Will it hurt? You will be asked to walk at your own pace, until you feel you can no longer continue due to symptoms. Pain or discomfort will not be greater than that of your everyday life. There is a small risk that you may stumble and fall on the treadmill. But, we will do everything we can to make sure that this does not happen.

Who will know? Personal records about this study will be kept confidential. Your data will be held in locked filing cabinets for 7 years in a safe area at the University of Alberta and can only be looked at by the research team. Any research data collected about you during this study will not identify you by name, only by your initials and a coded number. Any report published as a result of this study will not identify you by name. We may wish to look at the information again in the future. If so, it will first be looked at by a research ethics board.

Can you quit? You are free to withdraw from the research study at any time. Your continuing medical care will not be affected in any way. If any information gained from this or any other study becomes available which could influence your decision to continue in the study, you will be informed right away.

Will you be paid? You will be provided with parking coupons prior to each visit.

Contact Names and Telephone Numbers:
If you have concerns about your rights as a study participant, you may contact the Patient Relations Office of Capital Health, at 407-1040. If you have any concerns about this study, you may contact the Caritas Research Centre at (780) 930-5274. This office has no connection to the study investigators. If you have any questions, please do not hesitate to contact the persons listed below at any time.

Christy Tomkins (principal investigator)
Cell: 780.910.9318
Office: 780.492.1610

Email: ctomkins@ualberta.ca

Dr. Michele Crites Battie

Office: 780.492.5968

Email: mc.battie@ualberta.ca

Assessment of Walking in Lumbar Spinal Stenosis

PHYSICIAN FORM

Date: _____
Patient Name: _____
Patient Phone Number:_____

Does this patient have lumbar spinal stenosis, confirmed on imaging, with associated walking limitations or symptoms exacerbated with walking?

YES NO

Is this patient medically fit (no medical contraindications) to participate in this study, involving up to 30 minutes of self-paced walking?

YES NO

Has this patient provided verbal consent to be contacted by the research team to receive more information about the study, prior to consenting to study participation?

YES NO

Physician Name:_____

Physician Signature:_____

CONSENT TO BE CONTACTED FORM

I have read the information sheet, and agree to allow the investigators to contact me by phone to provide me with more information about the study. I authorize Dr._____ to release my contact information to Christy Tomkins for the purposes of being provided more information regarding the above-mentioned study. I am aware of the risks and benefits of consenting to be contacted, and that my consent may be revoked at any time.

NAME:_____
DATE:_____
PHONE NUMBER:_____

Signature:_____

Please return this form, along with the signed form given to you by the spine specialist to the research nurse or receptionist. Please keep the information sheet for your own records.

Thank you for your time and consideration!

Pre-Participation Screening Form

Potential Subject Name _____
Referring Physician _____
Phone Number _____

Date of Birth_____
Age_____

Hello, my name is Christy Tomkins and I am calling from the University of Alberta. I am calling to follow up on the information you received from _____last _____ regarding a study being conducted in Edmonton. We appreciate your willingness to consider participation in this study.

Persons with spinal stenosis sometimes experience increased symptoms or limitation with walking. Thus, walking capacity is very important in assessing spinal stenosis. The purpose of this study is to assess the validity of measures currently available for the assessment of walking capacity in people with lumbar spinal stenosis. We expect that the information from this study will help doctors, surgeons, and other health professionals in better evaluating walking, and treating individuals with lumbar spinal stenosis.

Your participation in this study will involve two or three visits of approximately 1 hour each to the University of Alberta Van Vliet Centre. You will be asked to fill out a series of short questionnaires regarding your back problem and your overall health. During your three visits you will also be asked to either walk around the track, or walk on a treadmill, at your own pace, for a maximum of 30 minutes, or until you feel you need to stop due to symptoms. You will walk on the track two times and the treadmill once. Greater detail regarding the study, is available on the information sheet provided to you by your spine specialist.

Do you have any questions regarding the study, or what will be asked of you?

Are you willing to participate in the study?

I will now ask you a series of questions designed to ensure that you are an appropriate candidate for the study, and to ensure your safety.

1. Are you limited in your ability to walk? YES NO

2. Is your walking limitation due to your back problem? YES NO

3. What limits your walking?
Pain ☐ Fatigue ☐ Numbness ☐ Shortness of breath ☐ Unbalance ☐

Other:_____

4a. Do you use any supportive walking devices? YES NO

b. If YES, what device?_____

5. Are you able to walk without your assistive device? YES NO

6. Do you have any conditions other than spinal stenosis which might limit your walking?
YES
NO

6b. If YES, please elaborate_____

_____

7. Have you previously had surgery for lumbar spinal stenosis? YES NO

7b. If YES, when?_____

8. Will you be having surgery for lumbar spinal stenosis? YES NO

8b If YES, when? _____

CONSENT FORM

Part 1 (to be completed by the Principal Investigator):

Title of Project: Assessment of Walking in Lumbar Spinal Stenosis

Principal Investigator(s): Christy Tomkins Phone Number(s): 492-1610

Co-Investigator: Michele Crites Battié Phone Number(s): 492-5968

Part 2 (to be completed by the research subject):

	Yes	No
Do you understand that you have been asked to be in a research study?	☐	☐
Have you read and received a copy of the attached Information Sheet?	☐	☐
Do you understand the benefits and risks involved in taking part in this research study?	☐	☐
Have you had an opportunity to ask questions and discuss this study?	☐	☐
Do you understand that you are free to withdraw from the study at any time, without having to give a reason and without affecting your future medical care?	☐	☐
Has the issue of confidentiality been explained to you?	☐	☐
Do you understand who will have access to your study records, including personally Identifiable health information?	☐	☐
Do you want the investigator(s) to inform your family doctor that you are participating in this research study? If so, give his/her name _____	☐	☐

Who explained this study to you? _____

I agree to take part in this study: YES ☐ NO ☐

Signature of Research Subject _____

(Printed Name) _____

Date: _____

Signature of Witness _____

I believe that the person signing this form understands what is involved in the study and voluntarily agrees to participate.

Signature of Investigator or Designee _____ Date _____

THE INFORMATION SHEET MUST BE ATTACHED TO THIS CONSENT FORM AND A COPY GIVEN TO THE RESEARCH SUBJECT

HISTORY QUESTIONNAIRE

Date: _____

Subject ID #:_____
Initials:_____

Age:_____

Male:_____ Female:_____

Address:_____

City:_____ Postal Code:_____

Home Phone:_____ Work/Cell Phone:_____

Emergency Contact Name and Phone:_____

Medical Information

Are you presently receiving treatment for a medical condition? YES NO
If YES, please explain _____
_____

Are you currently on any medications? YES NO
If YES, please indicate medications:_____
_____

Did you take any of these medications today? YES NO
Which medications did you take today?_____

Now a few questions about your back condition and walking ability:

Do you have back pain? ☐1 Yes ☐0 No

When did your back pain begin? (month/year):_____

Do you have leg pain? ☐1 Yes ☐0 No

Right Leg___ Left Leg___ Both_____

When did your leg pain begin? (month/year):_____

Walking Questions

How would you say your walking capacity is today, on a scale of 0 to 10, with 0 being your worst day, and 10 being your best day?:_____

If you were to go for a walk today, how far would you be able to walk at your own pace, on level ground before being forced to stop due to symptoms of lumbar spinal stenosis? _____m/km

Walking difficulty

By drawing a mark through the line as shown in this example, please indicate how much difficulty walking you have had during the past week:

Sample

None ——————————|——————————Maximum

Difficulty in walking

None ——————————————————————————Maximum

PHYSICAL FUNCTION SCALE OF THE SWISS SPINAL STENOSIS QUESTIONNAIRE

In the last <u>month</u> on a typical day:

1. How far have you been able to walk?

☐1 Over 2 miles

☐2 Over 2 blocks, but less than 2 miles

☐3 Over 50 ft., but less than 2 blocks

☐4 Less than 50 ft

2. Have you taken walks outdoors or in malls for pleasure?

☐1 Yes comfortably

☐2 Yes, but sometimes with pain

☐3 Yes, but always with pain

☐4 No

3. Have you been shopping for groceries or other items?

☐1 Yes, comfortably

☐2 Yes, but sometimes with pain

☐3 Yes, but always with pain

☐4 No

4. Have you walked around the different rooms in your house or apartment?

☐1 Yes comfortably

☐2 Yes, but sometimes with pain

☐3 Yes, but always with pain

☐4 No

5. Have you walked from your bedroom to the bathroom?

☐1 Yes, comfortably

☐2 Yes, but sometimes with pain

☐3 Yes, but always with pain

☐4 No

OSWESTRY DISABILITY INDEX

This is a 9-item questionnaire designed to give us information as to how your back or leg pain is affecting your ability to manage in everyday life. Please indicate which response best applies to you or which most clearly describes your problem **today.**

Section 1 - Pain Intensity

☐1 I have no pain at the moment
☐2 The pain is very mild at the moment
☐3 The pain is moderate at the moment
☐4 The pain is fairly severe at the moment
☐5 The pain is very severe at the moment
☐6 The pain is the worst imaginable at the moment

Section 2 - Personal Care (washing, dressing etc)

☐1 I can look after myself normally without causing extra pain
☐2 I can look after myself normally but it causes extra pain
☐3 It is painful to look after myself and I am slow and careful
☐4 I need some help but can manage most of my personal care
☐5 I need help every day in most aspects of self care
☐6 I do not get dressed, wash with difficulty and stay in bed

Section 3 - Lifting

☐1 I can lift heavy weights without extra pain
☐2 I can lift heavy weights but it gives me extra pain
☐3 Pain prevents me lifting heavy weights off the floor but I can manage if they are conveniently placed eg. on a table
☐4 Pain prevents me lifting heavy weights but I can manage light to medium weights if they are conveniently positioned
☐5 I can only lift very light weights

Section 4 – Walking*

☐1 Pain does not prevent me walking any distance
☐2 Pain prevents me from walking more than 2 kilometers
☐3 Pain prevents me from walking more than 1 kilometer
☐4 Pain prevents me from walking more than 500 meters
☐5 I can only walk using a stick or crutches
☐6 I am in bed most of the time

Section 5 - Sitting

☐1 I can sit in any chair as long as I like
☐2 I can only sit in my favorite chair as long as I like
☐3 Pain prevents me sitting more than one hour
☐4 Pain prevents me from sitting more than 30 minutes
☐5 Pain prevents me from sitting more than 10 minutes
☐6 Pain prevents me from sitting at all

Section 6 - Standing

☐1 I can stand as long as I want without extra pain
☐2 I can stand as long as I want but it gives me extra pain
☐3 Pain prevents me from standing for more than 1 hour
☐4 Pain prevents me from standing for more than 30 minutes
☐5 Pain prevents me from standing for more than 10 minutes
☐6 Pain prevents me from standing at all

Section 7 - Sleeping

☐1 My sleep is never disturbed by pain
☐2 My sleep is occasionally disturbed by pain
☐3 Because of pain I have less than 6 hours sleep
☐4 Because of pain I have less than 4 hours sleep
☐5 Because of pain I have less than 2 hours sleep
☐6 Pain prevents me from sleeping at all

Section 8 - Social Life

☐1 My social life is normal and gives me no extra pain
☐2 My social life is normal but increases the degree of pain
☐3 Pain has no significant effect on my social life apart from limiting my more energetic interests e.g. sport
☐4 Pain has restricted my social life and I do not go out as often
☐5 Pain has restricted my social life to my home
☐6 I have no social life because of pain

Section 9 - Travelling

☐1 I can travel anywhere without pain
☐2 I can travel anywhere but it gives me extra pain
☐3 Pain is bad but I manage journeys over two hours
☐4 Pain restricts me to journeys of less than one hour
☐5 Pain restricts me to short necessary journeys under 30 minutes
☐6 Pain prevents me from travelling except to receive treatment

HEALTH UTILITIES INDEX MARK 2-3

When answering the questions below, please think about your health and your ability to do things on a day-to-day basis, during the past week. Please select one answer that best describes your level of ability or disability **during the past week.**

1. Which one of the following best describes your ability, during the past week, to see well enough to read ordinary newsprint?
   - ☐1 Able to see well enough without glasses or contact lenses.
   - ☐2 Able to see well enough with glasses or contact lenses.
   - ☐3 Unable to see well enough even with glasses or contact lenses.
   - ☐4 Unable to see at all.

2. Which one of the following best describes your ability, during the past week, to see well enough to recognize a friend on the other side of the street?
   - ☐1 Able to see well enough without glasses or contact lenses.
   - ☐2 Able to see well enough with glasses or contact lenses.
   - ☐3 Unable to see well enough even with glasses or contact lenses.
   - ☐4 Unable to see at all.

3. Which one of the following best describes your ability, during the past week, to hear what was said in a group conversation with at least three other people?
   - ☐1 Able to hear what was said without a hearing aid.
   - ☐2 Able to hear what was said with a hearing aid.
   - ☐3 Unable to hear what was said even with a hearing aid.
   - ☐4 Unable to hear what was said but did not wear a hearing aid.
   - ☐5 Unable to hear at all.

4. Which one of the following best describes your ability, during the past week, to hear what was said in a conversation with one other person in a quiet room?
   - ☐1 Able to hear what was said without a hearing aid.
   - ☐2 Able to hear what was said with a hearing aid.
   - ☐3 Unable to hear what was said even with a hearing aid.
   - ☐4 Unable to hear what was said but did not wear a hearing aid.
   - ☐5 Unable to hear at all.

5. Which one of the following best describes your ability, during the past week, to be understood when speaking your own language with people who do not know you?
   - ☐1 Able to be understood completely.
   - ☐2 Able to be understood partially.
   - ☐3 Unable to be understood.
   - ☐4 Unable to speak at all.

6. Which one of the following best describes your ability, during the past week, to be understood when speaking your own language with people who know you well?

- ☐1 Able to be understood completely.
- ☐2 Able to be understood partially.
- ☐3 Unable to be understood.
- ☐4 Unable to speak at all.

7. Which one of the following best describes how you have been feeling during the past week?

- ☐1 Happy and interested in life.
- ☐2 Somewhat happy
- ☐3 Somewhat unhappy
- ☐4 Very unhappy
- ☐5 So unhappy that life was not worthwhile.

8. Which one of the following best describes the pain and discomfort you have experienced during the past week?

- ☐1 Free of pain and discomfort
- ☐2 Mild to moderate pain or discomfort that prevented no activities.
- ☐3 Moderate pain or discomfort that prevented a few activities.
- ☐4 Moderate to severe pain or discomfort that prevented some activities
- ☐5 Severe pain or discomfort that prevented most activities.

9. Which one of the following best describes your ability, during the past week, to walk? Note: Walking equipment refers to mechanical supports such as braces, a cane, crutches or a walker.

- ☐1 Able to walk around the neighborhood without difficulty, and without walking equipment.
- ☐2 Able to walk around the neighborhood with difficulty; but did not require walking equipment or the help of another person.
- ☐3 Able to walk around the neighborhood with walking equipment but without the help of another person.
- ☐4 Able to walk only short distances with walking equipment and required a wheelchair to get around the neighborhood.
- ☐5 Unable to walk alone, even with walking equipment. Able to walk short distances with the help of another person, and required a wheelchair to get around the neighborhood.
- ☐6 Unable to walk at all.

10. Which one of the following best describes your ability, during the past week, to use your hands and fingers? Note: Special tools refers to hooks for buttoning clothes, gripping devices for opening jars or lifting small items, and other devices to compensate for limitations of hands or fingers.

☐1 Full use of two hands and ten fingers.

☐2 Limitations in the use of hands or fingers, but did not require special tools or the help of another person.

☐3 Limitations in the use of hands or fingers, independent with use of special tools (did not require the help of another person).

☐4 Limitations in the use of hands or fingers, required the help of another person for some tasks (not independent even with use of special tools).

☐5 Limitations in the use of hands or fingers, required the help of another person for most tasks (not independent even with use of special tools).

☐6 Limitations in the use of hands or fingers, required the help of another person for all tasks (not independent even with use of special tools).

11. Which one of the following best describes your ability, during the past week, to remember things.

☐1 Able to remember most things

☐2 Somewhat forgetful

☐3 Very forgetful

☐4 Unable to remember anything at all.

12. Which one of the following best describes your ability, during the past week, to think and solve day to day problems?

☐1 Able to think clearly and solve day to day problems.

☐2 Had a little difficulty when trying to think and solve day to day problems.

☐3 Had some difficulty when trying to think and solve day to day problems.

☐4 Had great difficulty when trying to think and solve day to day problems.

☐5 Unable to think or solve day to day problems.

13. Which one of the following best describes your ability, during the past week, to perform basic activities?

☐1 Eat, bathe, dress and use the toilet normally.

☐2 Eat, bathe, dress and use the toilet independently with difficulty.

☐3 Required mechanical equipment to eat, bathe, dress or use the toilet independently.

☐4 Required the help of another person to eat, bathe, dress or use the toilet.

14. Which one of the following best describes how you have been feeling during the past week?

☐1 Generally happy and free from worry

☐2 Occasionally fretful, angry, irritable, anxious or depressed.

☐3 Often fretful, angry, irritable, anxious or depressed.

☐4 Almost always fretful, angry, irritable, anxious or depressed

☐5 Extremely fretful, angry, irritable, anxious or depressed; to the point of needing professional help.

15. Which one of the following best describes the pain or discomfort you have experienced during the past week?

☐1 Free of pain and discomfort

☐2 Occasional pain or discomfort. Discomfort relieved by non-prescription drugs or self-control activity without disruption of normal activities.

☐3 Frequent pain or discomfort. Discomfort relieved by oral medicines with occasional disruption of normal activities.

☐4 Frequent pain or discomfort; frequent disruption of normal activities. Discomfort required prescription narcotics for relief.

☐5 Severe pain or discomfort. Pain not relieved by drugs and constantly disrupted normal activities.

16. Overall, how would you rate your health during the past week?

☐1 Excellent

☐2 Very Good

☐3 Good

☐4 Fair

☐5 Poor

☐-9 No Response

CHANGE QUESTIONS
(Second and third visits only)

1. How would you say your walking capacity is today, compared to your last visit_____ days ago?

7__A great deal better
6__Moderately better
5__A little bit better
4__About the same
3__A little bit worse
2__Moderately worse
1__A great deal worse

2. How would you say your stenosis condition is today, compared with your last visit____days ago?

7__A great deal better
6__Moderately better
5__A little bit better
4__About the same
3__A little bit worse
2__Moderately worse
1__A great deal worse

5

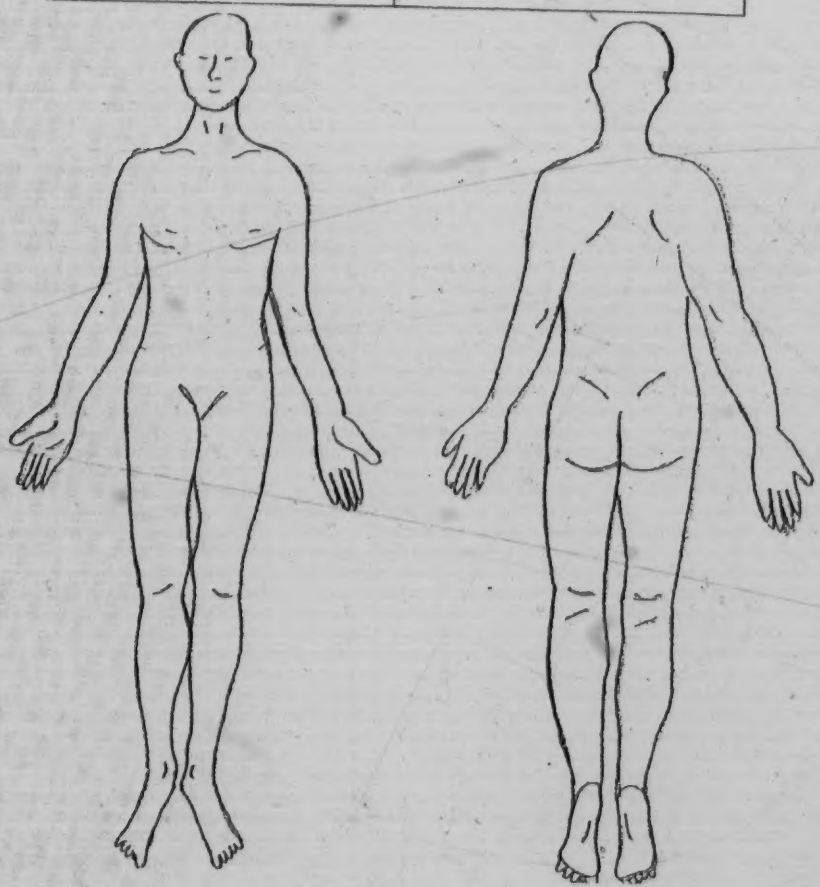
ID#_____
Initials_____
Date:_____

Pre-test	
Post-test	

BODY DIAGRAM

Please use the symbols below on the drawings to indicate the location of your symptoms.

Pain XXXX	Weakness ≡
Numbness OOOO	Tingling |||||



ID#_____
Initials_____
Date:_____

Pre-test	
Post-test	

Please place a mark through the line below to indicate your current pain level, on a scale from 'no pain' to 'the worst imaginable pain' for your back, right leg, and left leg.

Back

No pain _____ Worst imaginable

Right leg

No pain _____ Worst imaginable

Left leg

No pain _____ Worst imaginable

Self-paced Walking Test Data Collection Form

Subject ID #_____
Initials_____
Date_____
Age_____
Resting HR_____

Minute	Heart Rate
Before test/rest	
3	
6	
9	
12	
15	
18	
21	
24	
27	
Test termination	

Distance to first symptoms (m)	
Time to first symptoms (min)	
Total time (seconds)	
Number of laps	
Total distance (feet)	
Mean speed (m/min)	
Reason for test termination	__Symptoms of LSS __Fatigue/shortness of breath/dizziness __Pain or discomfort due to co-morbidities __Other:_____

First Symptoms

Back	Right leg	Left leg	Both legs	Buttocks
__Pain	__Pain	__Pain	__Pain	__Pain
__Numbness/ tingling	__Numbness/ tingling	__Numbness/ tingling	__Numbness/ tingling	__Numbness/ tingling
__Weakness	__Weakness	__Weakness	__Weakness	__Weakness
__Fatigue	__Fatigue	__Fatigue	__Fatigue	__Fatigue
__Other	__Other	__Other	__Other	__Other

Treadmill Test Data Collection Form

Subject ID #_____
Initials_____
Date_____
Age and DOB_____
Resting HR_____

Self-selected speed (mph) _____
Self-selected speed (m/min) _____

Minute	Heart Rate
Before test/rest	
3	
6	
9	
12	
15	
18	
21	
24	
27	
Test termination	

Distance to first symptoms (m)	
Time to first symptoms (min)	
Total time (seconds)	
Total distance	
Reason for test termination	__Symptoms of LSS __Fatigue/shortness of breath/dizziness __Pain or discomfort due to co-morbidities __Other:_____

First Symptoms

Back	Right leg	Left leg	Both legs	Buttocks
__Pain __Numbness/ tingling __Weakness __Fatigue __Other	__Pain __Numbness/ tingling __Weakness __Fatigue __Other	__Pain __Numbness/ tingling __Weakness __Fatigue __Other	__Pain __Numbness/ tingling __Weakness __Fatigue __Other	__Pain __Numbness/ tingling __Weakness __Fatigue __Other

APPENDIX B
DETAILED METHODOLOGY REGARDING DATA COLLECTION FOR STUDIES 2 AND 3

Subjects

Inclusion criteria: Central or combination (central + foraminal or lateral recess) lumbar spinal stenosis with associated walking limitations or symptoms exacerbated by walking. Subjects will be >45 years of age, and will have central or combination LSS confirmed by MRI/CT imaging, and by a spine specialist surgeon.

Exclusionary criteria: Foraminal/lateral recess stenosis without canal stenosis, peripheral vascular disease, severe cardiopulmonary, orthopaedic, or musculoskeletal conditions that would limit exercise/walking capacity, or make exercise medically inadvisable, as indicated by the spine specialist.

Subject Recruitment Procedure

Five spine specialist surgeons in Edmonton were briefed on the project at a meeting prior to subject recruitment. A set of forms was attached to the file of each patient who visited the clinics by a research nurse or administrative assistant, including a Physician Form, patient Information Sheet, and Consent to be Contacted Form. The spine specialists identified subjects who presented with lumbar spinal stenosis and had associated walking limitations or symptoms exacerbated with walking. At this point, physicians indicated on the Physician Form if they approved of the patient's involvement in the study, based on the protocols involving up to 30 minutes of self-paced walking. Thus, the surgeons medically screened subjects for participation.

Once a patient was identified as a potential subject, the physician signed the Physician Form and provided the patient with the Information and Consent to be Contacted forms. Patients will return the Physician Form as well as the completed Consent to be Contacted form to the nurse or administrative assistant, and keep one copy of the information form for their own records. The forms were then picked up as soon as possible from the clinics.

Pre-participation Screening

Subjects who consented to be contacted were called by me, as soon as possible after returning the Consent to be Contacted form. They were given an overview of the project, using the Pre-participation Screening form. Subjects were then asked a series of questions about their condition and their walking capacity, intended to ensure that they are appropriate for the study.

Inclusion criteria: Subjects had to answer YES to all of the following questions to be included in the study:

1. Are you limited in your ability to walk?
2. Is your walking limitation due to your back problem?
5. Are you able to walk without an assistive device?

Exclusion criteria: If subjects answered YES to the following question, they were excluded from the study:

1. Do you have any conditions other than spinal stenosis which might limit your walking?

If a subject indicated that they were having surgery for LSS within a month of the screening, or if they had had surgery within the last 12 months, they were excluded (given that the condition would be expected to change). The following questions were used to assess surgery-based exclusion criteria:

7. Have you previously had surgery for lumbar spinal stenosis?
7b. If YES, when was your surgery?
8. Will you be having surgery for lumbar spinal stenosis?
8b. If YES, when will your surgery be?

The remaining questions were used to collect descriptive data:

4. What limits your walking?
Pain ☐ Fatigue ☐ Numbness ☐ Shortness of breath ☐ Unbalance ☐

Other:_____

5a. Do you use any supportive walking devices?

5b. If YES to 5a., please elaborate_____

If it was decided that a subject was appropriate for the study, an appointment time for the first visit was be set, within the month following the pre-screening.

Data Acquisition

All data collection was conducted by the author. Subjects were asked to come to the Van Vliet Centre three times, with 7 days between each visit. Each visit was approximately 1 hour. Over the course of the study the total time commitment was be approximately 2.5-4 hours. Over the three visits subjects were asked to do the Self-Paced Walking Test twice and the treadmill protocol once. The order of the first two visits was randomized so that subjects had either the Self-Paced Walking Test or the treadmill protocol first, and the other test during the second visit. Subjects always completed the second Self-Paced Walking Test during the third visit.

Subjects were asked to wear comfortable clothing, and low-heeled shoes in which they would walk normally. Subjects were encouraged to wear the same shoes for all visits. If subjects were taking medications, they were asked to be consistent with their medications on the testing days (take the same medications on the days of visit one, two and three). In addition, they were asked to do their best to stay consistent with any other treatments they were receiving between visits. If at all possible the three tests were conducted at the same time of day.

Subjects were met at the parking lot of the Van Vliet Centre and walked to the track viewing area of the 'Butterdome'. On arriving at the track area subjects were informed that the straight-away on the track is 60m (~200ft), in order to provide the same distance reference to all subjects during the completion of the walking distance questions. Subjects were then guided to either the track, or the Work Physiology lab. At this time during the first visit the Information Sheet and Consent Form were reviewed with them. At this time and questions about the study were answered and the Consent Form was signed. Once subjects signed the Consent form they were asked to put on a heart rate monitor. Before either the track or treadmill test, subjects then sat for a minimum of 15 minutes. During the 15 minutes subjects completed the history and questionnaires

Upon completion of the appropriate forms, resting heart rate measurements were taken. Immediately prior to either the track walking or treadmill test, subjects were asked to record the location, nature (pain, numbness, weakness), and severity (mild, moderate, severe, very severe) of pre-test symptoms on a body diagram for descriptive purposes. Subjects also marked pre-test pain level on a 100mm 11 point visual analog pain scale for low back, left leg, and right leg.

Self-Paced Walking Test Protocol: Subjects were instructed to walk continuously at their own pace around the outer lane of the track, until they felt they had to stop due to symptoms of LSS (or other reasons), or until the time limit of 30 minutes was reached. At no time were subjects encouraged or prompted to continue. Subjects were also given a marker, which they were asked to raise in the air, and drop on the track at the point when they first experienced symptoms. At this time they were asked indicate the nature of the symptoms. If symptoms were present at the onset of the test, they were asked to drop the marker when they experienced a significant increase in symptoms. Test termination was defined as a complete stop of 3 seconds or more.

Throughout the test, the administrator followed at a comfortable distance, with a rolling wheel instrument to measure distance (Lufkin® Pro-Series Model PSMW38) and a stopwatch. The following information was collected during SPWT: Time/distance to onset of symptoms (time/distance=0 will be recorded if symptoms are present pre-test), time/distance to increase in symptoms if symptoms are present pre-test, nature of symptoms, total walking time, total distance walked, average walking speed, and the reason for test termination should they not walk for the full 30 minutes. Categories for early test termination included: LSS symptoms, general fatigue or shortness of breath/dizziness, and pain or discomfort due to co-morbidities. At the time of tests completion, subjects were asked to record the location, nature and severity of post-test symptoms on a body diagram and to complete a post-test VAS pain scale (back, left leg, right leg). Subjects were asked to wear the heart-rate monitor throughout testing. Heart rate was recorded every 3 minutes throughout the test, and the administrator looked for irregularities, such as rapid and large increases or decreases in heart rate. If such irregularities were noted, the test was stopped.

The distance wheel was checked monthly throughout the study for accuracy and reproducibility. During each monthly check the distance measured with the wheel was compared to a known distance on the track. This test was then repeated. Each check was successful with the wheel measuring the exact same distance as the known segment on the track for both trials.

Self-Paced Walking Test Script

Please walk at your own pace around the outer-most lane of the track. Continue to walk for as long as you can until you feel you have to stop due to symptoms of lumbar spinal stenosis, or for other reasons. I am also going to give you this marker. If you are in no pain or discomfort right now, I am going to ask you to raise the marker in the air, and drop it on the track at the point where you first experience pain or discomfort, and continue walking. I will also ask you to describe your symptoms to me at that time. If you are already in pain or discomfort, please raise the marker in the air and drop it when your symptoms increase noticeably. I will be following behind you, to make sure you are OK, and to measure the distance you walk. Do not over-exert yourself. This is not intended to be an exercise stress test, but a measurement of how far you can walk at your own pace. So again, continue to walk for as long as you can until you feel that you have to stop due to symptoms of LSS or for other reasons. Once you come to a full stop for more than 3 seconds, the test will be over. I will be asking you to indicate to me the reason why you stopped walking. Possible reasons for stopping include symptoms of LSS (pain, numbness, weakness, tingling), fatigue, shortness of breath, dizziness, or pain and discomfort due to other conditions. Do you have any questions?

Treadmill Protocol

For the treadmill test, subjects were asked to walk on a treadmill at 0% grade, until they felt they had to stop due to symptoms of LSS (or other reasons), or until the time limit of 30 minutes had been reached. At no time were subjects encouraged or prompted to continue. The protocol began with subjects standing on the treadmill. Subjects were asked to avoid holding the hand rails, and to maintain an upright posture. Subjects were permitted to place their hand, palm side down under the rail if it made them more comfortable or to steady themselves. The administrator slowly increased the speed to 1.2 miles per hour (the treadmill is set to read in mph). As soon as the treadmill started moving, time and distance were being recorded. At this time subjects were allowed to either increase or decrease the speed, to find a speed that was comfortable for them. Throughout the test subjects could ask the administrator to increase or decrease the speed of the treadmill to make walking more comfortable. Subjects were asked to notify the administrator when they experienced first symptoms, and to indicate the nature of the symptoms. If symptoms were present at the onset of the test, they were asked to notify the administrator when they experienced a significant increase in symptoms. The test was terminated at 30 minutes or by the subject when he or she felt they needed to stop due to symptoms of LSS, or for other reasons. The following information was collected during testing: Time to onset of symptoms (time=0 will be recorded if symptoms are present pre-test), time to increase in symptoms if symptoms are present pre-test, nature of symptoms, total walking time, total distance walked, and the reason for test termination should they not walk for the full 30 minutes. Categories for early test termination included: LSS symptoms, general fatigue or shortness of breath/dizziness, and pain or discomfort due to co-morbidities. At the time of test completion, subjects were asked to record the location, nature and severity of post-test symptoms on a body diagram and complete a post-test VAS pain scale (back, left leg, right leg). Heart rate was recorded every 3 minutes throughout the test, and the administrator looked for irregularities in the heart rate, such as rapid and large increases or decreases in heart rate. If such irregularities were noted, the test was stopped.

Treadmill Walking Test Script

Please step up onto the treadmill and make yourself comfortable. I am going to start the treadmill moving at a slow speed. Once the treadmill has reached 1.2mph I am going to ask you to use the arrows to increase or decrease the speed, until you are walking at a comfortable pace. Try to set the speed close to the pace you would normally walk. Once the speed has been set, you will continue to walk, with an upright posture. Avoid holding the handrails or leaning forward. If you need some stability, you may place your hands under the rails, but do not hold onto them. Throughout the test you may ask me to increase or decrease the speed of the treadmill to make walking more comfortable for you. Continue to walk for as long as you can until you feel you have to stop due to symptoms of lumbar spinal stenosis, or for other reasons. I will also ask you to indicate to me when you begin to have pain or discomfort. At that time I will ask you to describe your

symptoms. If you are already in pain or discomfort, please let me know when your symptoms increase noticeably. Remember, do not over-exert yourself. This is not intended to be an exercise test, but a measurement of how far you can walk at your own pace. So again, continue to walk for as long as you can until you feel that you have to stop due to symptoms of LSS or for other reasons. Once you feel you have to stop, press the red STOP button, and the treadmill will come to a gradual stop. This will be the end of the test. I will be asking you to indicate to me the reason why you stopped walking. Possible reasons for stopping include symptoms of LSS (pain, numbness, weakness, tingling), fatigue, shortness of breath, dizziness, or pain and discomfort due to other conditions. I will be standing right beside you throughout the test to ensure your safety. Do you have any questions?